Characterization of Fluorescent 8-Aryl-Guanine Mimics: From Mutagenic Lesions to Probes for Aptasensor Development

by

Michael Sproviero

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctorate of Philosophy
in
Chemistry and Toxicology

Guelph, Ontario, Canada
© Michael Sproviero, May, 2014
ABSTRACT

CHARACTERIZATION OF FLUORESCENT 8-ARYL-GUANINE MIMICS: FROM MUTAGENIC LESIONS TO PROBES FOR APTASENSOR DEVELOPMENT

Michael Sproviero
University of Guelph, 2014

Advisor: Professor R.A. Manderville

The genome of eukaryotic cells is constantly under attack. The DNA damage products resulting from such attacks are generally termed lesions, which can be generated by endogenous cellular processes or exogenous agents. Determining the structural impact these lesions have on DNA is critical for elucidating the mechanisms by which they exert their toxicity. Specifically, it is known that aryl radical species react with the C8-site of deoxyguanosine (dG) to form C8-aryl-dG adducts. Similar bulky DNA adducts may be formed through the covalent attachment of electrophilic species to the C8-site of dG. Typically, this leads to alteration of the overall structure and context of DNA, and such changes may lead to the propagation of a modified genetic code. Many of these adducts possess conformational heterogeneity within DNA, which has been proposed to account for various mutagenic outcomes.

In order to study the impact of C8-aryl-dG lesions, authentic oligonucleotide standards are required. Standard solid-phase DNA synthetic procedures are inefficient at generating C8-aryl-dG modified oligonucleotides. This is due to the inherent acid sensitivity of these adducts, which results in larger rates of hydrolysis. Thus, it was necessary to develop an efficient solid phase synthesis of oligonucleotides containing acid sensitive C8-aryl-dG adducts. Such a method was here optimized and is presented in this thesis.

Aside from altering the stability of the glycosidic bonds, C8-aryl dG adducts also display fluorescence emission. These fluorescent nucleosides have been shown to be sensitive to their solvent environment, base stacking and H-bonding interactions. This has provided a diagnostic
handle within DNA, and allowed C8-aryl-dG adducts to function as “probing adducts”. Specifically, C8-aryl-dG adducts are here utilized in different systems to report on: (1) various duplex perturbations, which stem from adduct formation within the NarI recognition sequence, (2) duplex–quadruplex exchange reactions within the thrombin binding aptamer, a common strategy used in DNA-based diagnostics, and (3) G-quadruplex conformational heterogeneity within the human telomeric sequence, which has implications in cancer and aging. The work presented in this thesis establishes 8-aryl-dG adducts as effective fluorescent probes, and supports their use in molecular diagnostic systems aimed at studying various DNA structure-function relationships.
ACKNOWLEDGMENTS

I have spent nearly a decade studying at the University of Guelph. I am grateful to have been able to pursue my education and training within an environment that fosters bright minds that are passionate, thoughtful and considerate. I am also grateful for the numerous teachers that were dedicated to their craft of stimulating student interest and enthusiasm for learning.

Specific acknowledgements go out to the people that have shaped and sculpted the man I am today. First and foremost thanks to my parents Mario and Terry Sproviero, you have been an inspiration and role models from day one; you have instilled in me the most important lessons of life: motivation, dedication, devotion, loyalty and love. Without your support I would not be here, you are my best friends. Thanks to my siblings (Matt and Sarah) for being far more than kinship and for your unwavering support with the struggles of life that are present at every corner. Thanks to my grandparents Ben & Judy Copan and Ida & Frank Sproviero, for the smiles on your faces every time I talked to you and the size of your hearts which made everything feel special. I am very grateful for my family and have developed an unbreakable bond and appreciation for them; they have been the most influential force that has pushed me to succeed.

Thanks to my brothers from other mothers: Tom Gibson, Kyle Stevens, Mike Kuska and Aaron Witham. It is quite rare that someone can spend as much time together as we have and not tire of one another. Tom and Kyle you have always provided me with the entertainment and joy I’ve needed away from work. Kuska we share so much more than a name, the weekly boys night was a time I will always remember. AW’s, in my math we have spent over 7200 hours (300 full days of time) together while we have been pursuing our PhD degrees; indeed we were quite lucky to have each other through the long lab nights and frustrations that research presents. I have learned so much from you and will always support the King of the North. I hope to add to our memories and friendship in coming years. I will miss our daily camaraderie but am interested to
see where our careers lead us and how we will continue to learn from one another despite being physically apart.

Last and certainly not least thanks to my advisor, the nucleic acid consigliere, Richard Manderville. You are an inspiration, thank you for your passion, effort, support and advice. Thank you for the many discussions we have shared on a daily basis over the last 5 years. I truly have cherished developing this research from the ground up with a man who genuinely loves what he does. You have been a great coach and together we have established meaningful, interesting research.
# TABLE OF CONTENTS

## CHAPTER 1: INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Background Information</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1. Historical Importance</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2. Primary/Secondary Structure of Nucleic Acids</td>
<td>3</td>
</tr>
<tr>
<td>1.1.3. Properties of DNA</td>
<td>4</td>
</tr>
<tr>
<td>1.1.3.1. Ionisation and Tautomerism</td>
<td>4</td>
</tr>
<tr>
<td>1.1.3.2. Hydrogen Bonding</td>
<td>5</td>
</tr>
<tr>
<td>1.1.3.3. Sugar-Phosphate Backbone and the Glycosyl Bond</td>
<td>8</td>
</tr>
<tr>
<td>1.1.3.4. Oxidative Properties</td>
<td>10</td>
</tr>
<tr>
<td>1.1.3.5. Photophysical Properties</td>
<td>10</td>
</tr>
<tr>
<td>1.1.4. DNA Tertiary Structure</td>
<td>16</td>
</tr>
<tr>
<td>1.1.4.1. Duplex DNA</td>
<td>16</td>
</tr>
<tr>
<td>1.1.4.2. Alternative DNA Structures</td>
<td>18</td>
</tr>
<tr>
<td>1.1.4.3. Overview of G4 DNA</td>
<td>18</td>
</tr>
<tr>
<td>1.1.5. Analysis of DNA Structure</td>
<td>20</td>
</tr>
<tr>
<td>1.1.5.1. Thermal Melting</td>
<td>20</td>
</tr>
<tr>
<td>1.1.5.2. Circular Dichroism</td>
<td>21</td>
</tr>
<tr>
<td>1.1.6. DNA Damage</td>
<td>26</td>
</tr>
<tr>
<td>1.1.6.1. Mutagenesis</td>
<td>26</td>
</tr>
<tr>
<td>1.1.6.2. Oxidative DNA damage</td>
<td>27</td>
</tr>
<tr>
<td>1.1.6.3. Hydrolysis of Nucleotides</td>
<td>28</td>
</tr>
<tr>
<td>1.1.6.4. Adduct Formation</td>
<td>29</td>
</tr>
<tr>
<td>1.1.6.5. C8-Aryl-Modified dG Nucleosides</td>
<td>30</td>
</tr>
<tr>
<td>1.1.6.6. Toxicological Relevance of C8 Adducts</td>
<td>33</td>
</tr>
<tr>
<td>1.2. Purpose of Research</td>
<td>35</td>
</tr>
<tr>
<td>1.3. References</td>
<td>38</td>
</tr>
</tbody>
</table>

## CHAPTER 2: SOLID PHASE SYNTHESIS OF OLIGONUCLEOTIDES CONTAINING ACID-SENSITIVE 8-ARYL-GUANINE ADDUCTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. BIG PICTURE:</td>
<td>46</td>
</tr>
<tr>
<td>2.2. General Introduction</td>
<td>47</td>
</tr>
<tr>
<td>2.2.1. Incorporation of Modified Nucleosides into DNA</td>
<td>54</td>
</tr>
<tr>
<td>2.2.1.1. Incorporation by the Phosphoramidite Approach</td>
<td>57</td>
</tr>
<tr>
<td>2.2.1.2. Automated DNA Synthesis</td>
<td>58</td>
</tr>
<tr>
<td>2.2.1.3. Interesting Modifications to Oligonucleotide Synthesis</td>
<td>61</td>
</tr>
<tr>
<td>2.3. Materials and Methods</td>
<td>65</td>
</tr>
<tr>
<td>2.3.1. General Methods</td>
<td>65</td>
</tr>
</tbody>
</table>
2.3.2. Synthesis of C8-Aryl-dG Phosphoramidites 65
  2.3.2.1. Suzuki-Miyaura Coupling of 8-Br-dG with Boronic Acids 68
    2.3.2.1.1. General Procedure 68
  2.3.2.2. N² Protection. 71
    2.3.2.2.1. General Procedure 71
  2.3.2.3. 5’OH DMT Protection 73
    2.3.2.3.1. General Procedure 74
  2.3.2.4. Alternative 5’OH Protection with DMPx 77
    2.3.2.4.1. General Procedure for DMPx Protection of 8-Aryl-dG Nucleosides 80
  2.3.2.5. Phosphoramidite Reaction 83
    2.3.2.5.1. General Procedure 84
  2.3.3. Synthesis of C8-Aryl-dG Modified Oligonucleotides 90
    2.3.3.1. Mass Spectrometry Analysis 91

2.4. Results and Discussion 93

2.5. Conclusions 102

2.6. References 103

CHAPTER 3: IMPACT OF 8-ARYL-GUANINE PROBES ON DNA DUPLEX STRUCTURE 107

3.1. BIG PICTURE 108

3.2. Background 109
  3.2.1. DNA Damage 109
  3.2.2. Determination of Structure-Activity Relationships 116
  3.2.3. Fluorescent Nucleobase Analogues 121
  3.2.4. C8-Aryl-dG Modifications as “Probing Adducts” 123

3.3. Materials and Methods 125
  3.3.1. General Methods 125
  3.3.2. Preparation of Oligonucleotide Substrates 125
  3.3.3. Computer Simulations 126

3.4. Results and Discussion 126
  3.4.1. Nucleoside Properties 126
  3.4.2. 8-Aryl-dG Adduct Impact on dsDNA Structure and Stability 132
    3.4.2.1. Thermal Melting Analysis 132
    3.4.2.2. Circular Dichroism 140
  3.4.3. Conformational preferences of adducts from MD simulations 144
  3.4.4. Fluorescence Response of 8-Aryl-dG Probing Adducts 146

3.5. Conclusions 154
CHAPTER 4: HARNESSING G-TETRAD SCAFFOLDS WITHIN G-QUADRUPLEX FORMING APTAMERS FOR FLUORESCENCE DETECTION STRATEGIES

4.1. BIG PICTURE

4.2. Background
   4.2.1. Generation of an Aptamer
   4.2.2. Chemical Modification of Aptamers
      4.2.2.1. Imparting Chemical Diversity
      4.2.2.2. Improving the Pharmacokinetic Profile of the Aptamer
      4.2.2.3. Aptasensors
   4.2.3. Thrombin Binding Aptamer

4.3. Materials and Methods
   4.3.1. Oligonucleotide Synthesis and Purification
   4.3.2. Mass Spectrometry Analysis

4.4. Results and Discussion
   4.4.1. Impact on Structure and Stability
      4.4.1.1. Thermal Melting Analysis
      4.4.1.2. Circular Dichroism
   4.4.2. Probe Performance
   4.4.3. Titrations
      4.4.3.1. K⁺ Titrations
      4.4.3.2. Thrombin Binding Titrations

4.5. Conclusions

4.6. References

CHAPTER 5: INCORPORATION OF FURDG WITHIN THE HUMAN TELOMERIC SEQUENCE

5.1. BIG PICTURE:

5.2. General Introduction
   5.2.1. Structure and Stability of Telomeres:
   5.2.2. Structural Polymorphism of H-Telo G-Quadruplexes
   5.2.3. Telomerase Therapeutics
      5.2.3.1. G-Quadruplex Stabilizing Molecules

5.3. Materials and Methods
   5.3.1. General Methods
5.3.2. Oligonucleotide Synthesis and Purification 238
5.3.3. Mass Spectrometry Analysis 239

5.4. Results and Discussion 239
5.4.1. Impact on Structure and Stability 239
  5.4.1.1. Thermal Melting Analysis 240
  5.4.1.2. Circular Dichroism 245
5.4.2. Fluorescence Response 249
5.4.3. Modulation of Folding Topologies 255
  5.4.3.1. Response to Water Activity 256
  5.4.3.2. Response to Conformational Selection 261
    5.4.3.2.1. Addition of PEG 261
    5.4.3.2.2. Addition of NMM 265

5.5. Conclusions 272

5.6. References 275

APPENDIX A: GENERAL EXPERIMENTAL METHODS 281

APPENDIX B: CHARACTERIZATION OF SYNTHESIZED PRODUCTS BY $^1$H, $^{13}$C, $^{31}$P NMR SPECTROSCOPY 289

APPENDIX C: CHARACTERIZATION OF C8-ARYL-G MODIFIED Oligonucleotides BY ESF MASS SPECTROMETRY AND UV SPECTROSCOPY 344
# LIST OF TABLES

| Table 1 - 1 | The pKa values for bases as nucleosides and nucleotides. | 5 |
| Table 1 - 2 | Photophysical parameters of nucleosides. | 10 |
| Table 1 - 3 | Sources of reactive oxygen species | 27 |
| Table 2 - 1 | ESI-MS Analysis of 8-aryl-dG Modified NarI(12) oligonucleotides. | 92 |
| Table 2 - 2 | ESI-MS Analysis of 8-aryl-dG Modified NarI(22) oligonucleotides | 93 |
| Table 2 - 3 | Half-lives ($t_{1/2}$) for hydrolysis of 8-aryl-dG adducts | 94 |
| Table 2 - 4 | Summary of first-order rate constants ($k_{obs}$) and half-lives ($t_{1/2}$) for hydrolysis of Fur-dG and its derivatives | 95 |
| Table 2 - 5 | Summary of yields for 8-aryl-dG modified NarI oligonucleotides | 100 |
| Table 3 - 1 | Solvatochromic properties of 8-aryl-dG adducts | 127 |
| Table 3 - 2 | Tabulated excitation and emission spectral information for 8-aryl-dG mNarI oligonucleotide duplexes. | 148 |
| Table 4 - 1 | Comparison between aptamers and antibodies | 168 |
|             | ESI-MS analysis of mTBA oligonucleotides | 176 |
| Table 4 - 2 | Tabulated $T_m$ data for C8-aryl-dG modified mTBA oligonucleotides. | 179 |
| Table 4 - 3 | Tabulated excitation/emission data for Fur-dG mTBA. | 196 |
| Table 4 - 4 | Tabulated excitation/emission data for Bth-dG mTBA. | 197 |
| Table 4 - 5 | Tabulated excitation/emission data for CNPb-dG mTBA. | 198 |
| Table 4 - 6 | Tabulated excitation/emission data for Q-dG mTBA. | 199 |
| Table 4 - 7 | Tabulated excitation/emission data for Py-dG mTBA. | 199 |
| Table 4 - 8 | Tabulated excitation/emission spectra of the dual modified mTBA oligonucleotides. | 204 |
| Table 4 - 9 | Tabulated association constants and corresponding coefficient of determination values for thrombin binding by 8-aryl-dG mTBA oligonucleotides | 214 |
| Table 4 - 10 | Tabulated association constants and corresponding coefficient of determination values for thrombin binding by 8-aryl-dG mTBA oligonucleotides | 214 |
| Table 5 - 1 | ESI-MS analysis of mHTelo22 oligonucleotides | 239 |
Table 5 - 2  Tabulated $T_m$ values for $^{15}$dG mHTelo oligonucleotides folding into duplex or quadruplex structures.  240

Table 5 - 3  Tabulated photophysical properties of mHTelo oligonucleotides.  252
| Figure 1 - 1 | Nucleoside/nucleotide composition and structures of the major purine and pyrimidine nucleobases of nucleic acids in their dominant tautomeric forms. |
| Figure 1 - 2 | H-bonding patterns of the Watson-Crick Base Pairs |
| Figure 1 - 3 | H-bonding patterns of Hoogsteen base pairing |
| Figure 1 - 4 | The covalent backbone of DNA, showing the linkage of successive nucleotide units by phosphodiester bonds. |
| Figure 1 - 5 | Torsional angles for a nucleotide within the DNA duplex. |
| Figure 1 - 6 | Conformational equilibrium between syn C3'-endo, and anti C2'-endo for 2'-dG. Figure also highlights the Watson-Crick and Hoogsteen H-bonding faces. |
| Figure 1 - 7 | Generic Jablonski diagram illustrating: A – Absorbance; IC – Internal Conversion; NRD – Non-Radiative Decay; IC – Intersystem Crossing; F – Fluorescence; P – Phosphorescence. |
| Figure 1 - 8 | Depiction of various external fluorescence modifications linked to the terminus of an oligonucleotide. (a) Fluorescein 6-FAM; (b) NHS-Rhodamine; (c) Cy3 dye. |
| Figure 1 - 9 | Depiction of various internal fluorescent analogues. (a) 2-AP; (b) dG Pteridine derivatives; (c) dA Pteridine derivatives; (d) – (f) polycyclic hydrocarbon based fluorescent nucleobase derivatives. dR corresponds to the 2’-deoxyribose sugar. |
| Figure 1 - 10 | Depiction of B-DNA. Hydrophilic phosphate backbones are colored blue while the hydrophobic purine and pyrimidine bases are colored red. |
| Figure 1 - 11 | Structure of a G-tetrad which can stack vertically to assemble G quadruplexes. |
| Figure 1 - 12 | Various strand alignments in the formation of G quadruplexes. Quadruplexes can be formed intra- or intermolecularly and can be comprised of single or multiple (up to four) oligonucleotide strands |
| Figure 1 - 13 | Characteristic CD spectra for B-, Z-, and A-form DNA. Reproduced from: http://homepage.univie.ac.at/johannes.winkler/cd.html |
| Figure 1 - 14 | CD spectra of guanine quadruplexes. Left side: the parallel stranded quadruplex [d(G₄)]₂ stabilized by 16 mM K⁺; right side: Na⁺-induced anti-parallel bimolecular quadruplex of [d(G₄T₄G₄)]₂. The triangles in the sketches indicate guanines and point in the 5’–3’ direction. Reproduced from reference with permission. |
| Figure 1 - 15 | The G-Quartet (a) shown on its “head” (H) side (from donor to acceptor H-bonding runs clockwise); (b) The “tail” side (donor from acceptor H-bonding runs counter clockwise); (c) Heteropolar stacking arrangement of two G-tetrad (head to tail stacked with one another); (d) Homopolar stacking arrangement of two G-tetrad (head to head stacked with one another). |
Figure 1 - 16  The structure of 8-oxo-G, guanidinohydantoin heterocycle (Gh) and spirominodihydantoin (Sp) lesions.

Figure 1 - 17  Site-specific modifications of dG by chemical carcinogens.

Figure 1 - 18  Structures of commonly studied N-linked dG DNA adducts.

Figure 1 - 19  Depictions of the three major conformers of C8-modified deoxyguanosine adducts within duplex DNA. Aryl is used to denote the adducted moiety.

Figure 2 - 1  Deglycosylation of substituted benzene C8-aryl adducts driven by the relief of steric strain upon sugar loss.

Figure 2 - 2  General synthesis of 8-aryl-dG adducts using 8-Br-dG and the Suzuki-Miyaura coupling methodology.

Figure 2 - 3  Methods for Synthesis of Oligonucleotides Containing DNA Adducts at Defined Positions.

Figure 2 - 4  General depiction of the unmodified nucleobase phosphoramidites. Squares are used to depict a 5'OH protecting group while the nucleobase exocyclic amines are protected with a group symbolized by triangles.

Figure 2 - 5  Attachment of a nucleoside to the CPG solid support for Solid Phase DNA synthesis.

Figure 2 - 6  Solid-phase two step phosphoramidite oligodeoxynucleotide synthesis cycle.

Figure 2 - 7  Structures of pixyl and 2,7-dimethyloxypixyl (DMPx) protecting groups.

Figure 2 - 8  Comparison between the DMT, Px and DMPx protecting groups utilized for nucleoside 5'OH protection.

Figure 2 - 9  Synthesis of DMPx protected nucleosides for utilization in solid phase synthesis.

Figure 2 - 10  Depicting the differences between the rates of detritylation for 5'-O-DMT-N2-dimethylformamidyl-8-(2''-furanyl)-2'-dG in (i) 3% DCA and (ii) 0.5% DCA in acetonitrile.

Figure 2 - 11  HPLC traces for solid phase DNA synthesis of thymidine decamer (5'-T7Fur dGTT) utilizing (top) DMT protected nucleosides; and (bottom) DMPx protected nucleosides.

Figure 2 - 12  HPLC trace for solid phase DNA synthesis of (5'-DMT-TFur dGTT) utilizing standard protocols.

Figure 2 - 13  HPLC trace for the solid phase DNA synthesis of cytidine decamer (5'-CC7Fur dGCC) using standard protocols.

Figure 2 - 14  HPLC traces for solid phase DNA synthesis of NarI(12) (5'-CTCGGC7Fur GCCATC) utilizing (top) DMT protected nucleosides; and (bottom)
DMPx protected nucleosides.

Figure 3 - 1 Structure of various C8-aryl-dG analogues. (a) Adduct formation from PAH carbocation formation; (b) adduct formation due to estroquinone radical formation (c) adduct formation from activation of OTA; (d) adduct formation from phenoxy radical, or aryl radical via reduction of arenediazonium ions.

Figure 3 - 2 Structure of various N2-guanine adducts utilized to determine the impact of adduct size on mutagenicity.

Figure 3 - 3 (a) AF-dG lesions exist in an anti-syn conformational equilibrium. (b) AAF-dG lesions exist primarily in the syn orientation. (c) Structure of various C8-N-acetyl-aryl-dG lesions studied.

Figure 3 - 4 Structure of various pteridine fluorescent nucleobase analogues.

Figure 3 - 5 How the family of 8-aryl-dG probing adducts (bottom) compare as structural analogues of various authentic DNA adducts (top).

Figure 3 - 6 Generation of the CT species upon excitation of CNPh-dG.

Figure 3 - 7 Depiction of the dual fluorescence observed for the push-pull QdG adduct. (a) Observed emission spectra for QdG in water (red trace), CH3CN (dashed blue trace) and CHCl3 (dotted green trace); (b) proposed emission in H2O; (c) proposed emission in CH3CN; (d) proposed emission in CHCl3.

Figure 3 - 8 Twist angles, dipole moments and ground-state B3LYP/6-31G(d) global minima orbital density plots of 8-aryl-dG adducts.

Figure 3 - 9 CD Spectra for mNarI duplexes. (a) Narl(X):Narl'(C) (b) Narl(X):Narl'(G) (c) Narl(X):Narl'(10) (d) Narl(X):Narl'(THF). Black dotted = dG; Black solid = PhdG; Orange = PydG; Red = FurdG; Green = BthdG; Blue = QdG; Purple = CNPh-dG.

Figure 3 - 10 CD spectral overlay of all PydG mNarI duplexes, for complementary strands purple = Narl'(THF), blue = Narl'(10), red = Narl'(C), green = Narl'(G).

Figure 3 - 11 Lowest energy conformations (within 15 kJ mol⁻¹ of free energy) for the studied adducts paired against cytosine. The relative free energies of two competing conformations of Fur-dG are provided in bold in parentheses.

Figure 3 - 12 Lowest energy conformations (within 15 kJ mol⁻¹ of free energy) for the studied adducts paired against guanine. The relative free energies of two competing conformations of QdG are provided in bold in parentheses.

Figure 3 - 13 Excitation and emission spectra of 8-aryl-dG mNarI oligonucleotide duplexes. (a) PhdG; (b) FurdG; (c) BthdG; (d) PydG; (e) CNPhdG; (f) QdG. Red = Narl'(C); Green = Narl'(THF); Blue = Narl'(G); Purple Narl'(10); Black ssDNA.

Figure 4 - 1 Structures of various modified nucleosides used in SELEX.

Figure 4 - 2 (a) Structure of a G-tetrad with M⁺ cation in central core; (b) TBA with syn-G's
colored red and *anti*-G’s colored blue; (c) Structure of the ternary Thrombin-TBA-Na⁺ complex; (d) Structure of the ternary Thrombin-TBA-K⁺ complex.

Figure 4 - 3  (Left) Depiction of mTBA highlighting positions of C8-aryl modification. (Right) Structure of C8-aryl-dG modifications utilized.

Figure 4 - 4  *T*ₘ profiles for 8-aryl-dG mTBA oligonucleotides. Solid blue line represents duplex *T*ₘ; solid red line represents K⁺ stabilized quadruplex *T*ₘ; dotted red line represents Na⁺ stabilized quadruplex *T*ₘ.

Figure 4 - 5  *T*ₘ profiles of various doubly modified mTBA oligonucleotides. (a) **CNPh**dG@8/Fur dG@5; (b) **CNh**dG@8/Fur dG@10; (c) Fur dG@1/Fur dG@8; (d) Fur dG@5/Fur dG@8. Solid red line corresponds to TBA-K⁺ quadruplex *T*ₘ; Dotted line is TBA-Na⁺ quadruplex *T*ₘ; Solid Blue line corresponds to TBA-Na⁺ duplex.

Figure 4 - 6  CD spectra of mTBA duplexes. (a) modification at position 5; (b) modification at position 6; (c) modification at position 8. Black = unmodified TBA; Orange = Fur dG mTBA; Red = Fur dG mTBA, Blue = CNPh dG mTBA; Purple = Q dG mTBA; Green = Bth dG mTBA.

Figure 4 - 7  CD spectra for mTBA oligonucleotide quadruplexes. Black = unmodified TBA; Orange = Fur dG mTBA; Red = Fur dG mTBA, Blue = CNPh dG mTBA; Purple = Q dG mTBA; Green = Bth dG mTBA.

Figure 4 - 8  Crystal structure of K⁺-TBA with G5 colored red and C8 of G5 highlighted in green. Adapted from previously reported crystal structure.

Figure 4 - 9  Crystal structure of K⁺-TBA with G8 colored red and C8 of G8 highlighted in green. Adapted from previously reported crystal structure.

Figure 4 - 10  Excitation and Emission spectra for mTBA oligonucleotides. Solid red line represents modification at position 5 in K⁺; dotted red line represents modification at position 5 in Na⁺; blue line represents modification at position 6; black line represents modification at position 8.

Figure 4 - 11  Excitation and emission spectra of (a) **Fur**dG@5/CNPh dG@8 (b) **Fur**dG@10/CNPh dG@8 (c) **Fur**dG@1/Fur dG@8; (d) **Fur**dG@5/Fur dG@8 mTBA oligonucleotides. Dotted line represents duplex. Dashed line represents Na⁺ quadruplex. Solid Line represents K⁺ quadruplex.

Figure 4 - 12  Probe fluorescence response in duplex–quadruplex exchange mediated by K⁺. Duplex samples represented by solid trace. KCl was then added in 5 mM aliquots (dotted traces). Inserts show CD overlays of initial duplex (solid trace) and G-quadruplex after the final KCl addition (dotted trace). (a) Titration of **Fur**dG@5 mTBA:10-mer with KCl at 25 °C, (b) titration of **Fur**dG@5/CNPh dG@8-mTBA:15-mer with KCl at 30°C.

Figure 4 - 13  ssDNA to quadruplex fluorescence titration of TBA **Fur**dG@5 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.
Figure 4 - 14  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@6 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 15  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@8 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 16  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Bth}dG@5 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 17  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Bth}dG@8 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 18  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{CNPr}dG@5 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 19  dsDNA to quadruplex fluorescence titration of TBA \textsuperscript{CNPr}dG@5 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 20  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@5/\textsuperscript{CNPr}dG@8 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 21  ssDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@5/\textsuperscript{Py}dG@8 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 22  dsDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@5 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 23  dsDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@5/\textsuperscript{CNPr}dG@8 with thrombin. Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.
Figure 4-24 dsDNA to quadruplex fluorescence titration of TBA $^{\text{Fur}}$dG@1/$^{\text{Py}}$dG@8 with thrombin. Here excitation and emission wavelengths were set to monitor the $^{\text{Py}}$dG moiety ($\lambda_{\text{ex}}$ 360 nm, $\lambda_{\text{em}}$ 460 nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4-25 dsDNA to quadruplex fluorescence titration of TBA $^{\text{Fur}}$dG@1/$^{\text{Py}}$dG@8. Here excitation and emission wavelengths were set to monitor the $^{\text{Fur}}$dG probe ($\lambda_{\text{ex}}$ 315 nm, $\lambda_{\text{em}}$ 380 nm).

Figure 4-26 ssDNA to fluorescence titration of TBA $^{\text{Fur}}$dG@5 with serum albumin. Insert: % Change in fluorescent signal vs. [bovine serum albumin].

Figure 4-27 ssDNA to fluorescence titration of TBA $^{\text{CNPh}}$dG@5 with bovine serum albumin. Insert: % Change in fluorescent signal vs. [bovine serum albumin].

Figure 5-1 Representation of a T loop structure

Figure 5-2 Various folding topologies of the human telomeric 22-mer sequence (HTel22). Reproduced from reference with permission. Grey rectangles represent dG’s in the anti conformation, red rectangles represent dG’s in the syn-orientation. Blue spheres represent dA yellow spheres represent dT.

Figure 5-3 Depiction of the reorganization of HTelo22 from the anti-parallel, basket type structure in Na$^+$ to the various hybrid forms in K$^+$. Reproduced from reference with permission.

Figure 5-4 (a) Possible model of duplex, i-motif and G-quadruplex association for the HTelo repeat, Reproduced from reference with permission. (b) The hybrid-type telomeric G-quadruplex structures can be readily folded and stacked end to end to form compact-stacking structures for multimers in the elongated telomeric DNA, reproduced from reference with permission. (c) Parallel-type telomeric G-quadruplex structures folding and stacked end to end with same direction or alternate direction stacking for connection between G-quadruplex blocks, reproduced from reference with permission.

Figure 5-5 Various G-quadruplex stabilizing molecules

Figure 5-6 Depiction of the various $T_m$ profiles obtained for mHTelo oligonucleotides. Blue trace refers to duplex formation while red and black traces refer to quadruplex formation.

Figure 5-7 Circular Dichroism spectra for mHTelo oligonucleotides. Black line depicts unmodified HTelo oligonucleotide while the red line depicts $^{\text{Fur}}$dG@3, blue Line depicts $^{\text{Fur}}$dG@8, and green line depicts $^{\text{Fur}}$dG@10. (a) Na$^+$ stabilized duplex; (b) Na$^+$ stabilized quadruplex; (c) K$^+$ stabilized quadruplex

Figure 5-8 Excitation and Emission spectra of $^{\text{Fur}}$dG within (a) duplex, (b) Na$^+$ stabilized quadruplexes, (c) K$^+$ stabilized quadruplexes. Red line corresponds to $^{\text{Fur}}$dG@3; blue line corresponds to $^{\text{Fur}}$dG@8; green line corresponds to $^{\text{Fur}}$dG@10.
Figure 5 - 9  CD spectral changes upon addition of CH$_3$CN to mHTelo oligonucleotides. Solid lines represent 100% K$^+$-buffer solutions while dotted lines are 42.5% v/v CH$_3$CN/Buffer. (a) Unmodified sequence (b) Fur$_3$G@3 (c) Fur$_8$G@8 (d) Fur$_{10}$G@10

Figure 5 - 10  Excitation and emission spectral overlays comparing various Fur$_3$G mHTelo oligonucleotides in K$^+$ buffer (solid line) or 42.5% v/v CH$_3$CN/K$^+$ buffer (dotted line). (a) Fur$_3$G@3; (b) Fur$_8$G@8; (c) Fur$_{10}$G@10.

Figure 5 - 11  CD spectral changes upon addition of PEG-600 to mHTelo oligonucleotides. Solid lines represent 100% K$^+$-buffer solutions while dotted lines are 42.5% v/v PEG-600/Buffer. (a) Unmodified sequence (b) Fur$_3$G@3 (c) Fur$_8$G@8 (d) Fur$_{10}$G@10.

Figure 5 - 12  Excitation and emission spectral overlays comparing Fur$_3$G mHTelo oligonucleotides in K$^+$ buffer (solid line) or 42.5% v/v PEG-600/K$^+$ buffer (dotted line). (a) Fur$_3$G@3; (b) Fur$_8$G@8; (c) Fur$_{10}$G@10.

Figure 5 - 13  CD spectral changes upon addition of NMM to mHTelo oligonucleotides. Solid lines represent 100% K$^+$-buffer solutions while dotted lines correspond to solutions containing 5 equivalents NMM. (a) Unmodified sequence (b) Fur$_3$G@3 (c) Fur$_8$G@8 (d) Fur$_{10}$G@10.

Figure 5 - 14  Excitation and emission spectral overlays comparing various Fur$_3$G mHTelo oligonucleotides in K$^+$ buffer (solid line) or K$^+$ buffer with 5 equivalents NMM (dotted line). (a) Fur$_3$G@3; (b) Fur$_8$G@8; (c) Fur$_{10}$G@10.

Figure 5 - 15  Spectral overlay depicting Fur$_3$G emission (blue) and NMM excitation (red).

Figure 5 - 16  Molecular orbital schematic for RET.

Figure 5 - 17  Addition of NMM to Fur$_3$G mHTelo quadruplexes. (a) Fur$_3$G@3 (b) Fur$_8$G@8 (c) Fur$_{10}$G@10. Prefolded quadruplex in 42.5% v/v CH$_3$CN/K$^+$ buffer (solid line) was subjected to 4.5 μM additions of NMM (dotted traces).

Figure B - 1  $^1$H NMR spectrum of 8-quinoline-boronic acid in CD$_3$OD.

Figure B - 2  $^{13}$C NMR spectrum of 8-quinoline-boronic acid in CD$_3$OD.

Figure B - 3  $^1$H NMR spectrum of 8-(2''-furanyl)-2'-dG (1a) in DMSO-d$_6$.

Figure B - 4  $^{13}$C NMR spectrum of 8-(2''-furanyl)-2'-dG (1a) in DMSO-d$_6$.

Figure B - 5  $^1$H NMR spectrum of 8-phenyl-2'-dG (1b) in DMSO-d$_6$.

Figure B - 6  $^{13}$C NMR spectrum of 8-phenyl-2'-dG (1b) in DMSO-d$_6$.

Figure B - 7  $^1$H NMR spectrum of 8-(4''-cyanophenyl)-2'-dG (1c) in DMSO-d$_6$.

Figure B - 8  $^{13}$C NMR spectrum of 8-(4''-cyanophenyl)-2'-dG (1c) in DMSO-d$_6$.

Figure B - 9  $^1$H NMR spectrum 8-(8''-quinolyl)-2'-dG (1d) in DMSO-d$_6$. 
Figure B - 10  $^{13}$C NMR spectrum 8-(8''-quinolyl)-2'-dG (1d) in DMSO-d$_6$.  
Figure B - 11  $^1$H NMR spectrum of 8-(2''-benzothienyl)-2'-dG (1e) in DMSO-d$_6$.  
Figure B - 12  $^{13}$C NMR spectrum of 8-(2''-benzothienyl)-2'-dG (1e) in DMSO-d$_6$.  
Figure B - 13  $^1$H NMR spectrum 8-(pyrenyl)-2'-dG (1f) in DMSO-d$_6$.  
Figure B - 14  $^{13}$C NMR spectrum 8-(pyrenyl)-2'-dG (1f) in DMSO-d$_6$.  
Figure B - 15  $^1$H NMR spectrum of N$^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (2a) in DMSO-d$_6$.  
Figure B - 16  $^{13}$C NMR spectrum of N$^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (2a) in DMSO-d$_6$.  
Figure B - 17  $^1$H NMR spectrum of N$^2$-(dimethylformamidyl)-8-phenyl-2'-dG (2b) in DMSO-d$_6$.  
Figure B - 18  $^{13}$C NMR spectrum of N$^2$-(dimethylformamidyl)-8-phenyl-2'-dG (2b) in DMSO-d$_6$.  
Figure B - 19  $^1$H NMR spectrum of N$^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (2c) in DMSO-d$_6$.  
Figure B - 20  $^{13}$C NMR spectrum of N$^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (2c) in DMSO-d$_6$.  
Figure B - 21  $^1$H NMR spectrum N$^2$-(dimethylformamidyl)-8-(8''-quinolyl)-2'-dG (2d) in DMSO-d$_6$.  
Figure B - 22  $^{13}$C NMR spectrum N$^2$-(Dimethylformamidyl)-8-(8''-Quinolyl)-2'-dG (2d) in DMSO-d$_6$.  
Figure B - 23  $^1$H NMR spectrum of N$^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (2e) in DMSO-d$_6$.  
Figure B - 24  $^{13}$C NMR spectrum ofN$^2$-(Dimethylformamidyl)-8-(2''-Benzothienyl)-2'-dG (2e) in DMSO-d$_6$.  
Figure B - 25  $^1$H NMR spectrum of N$^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (2f) in DMSO-d$_6$.  
Figure B - 26  $^{13}$C NMR spectrum N$^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (2f) in DMSO-d$_6$.  
Figure B - 27  $^1$H NMR spectrum of 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (3a) in CD$_3$CN.  
Figure B - 28  $^{13}$C NMR spectrum of 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (3a) in CD$_3$CN.
Figure B - 29 $^1$H NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (3b) in DMSO-d$_6$.

Figure B - 30 $^{13}$C NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (3b) in DMSO-d$_6$.

Figure B - 31 $^1$H NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (3c) in CD$_2$Cl$_2$.

Figure B - 32 $^{13}$C NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dGe (3e) in CD$_2$Cl$_2$.

Figure B - 33 $^1$H NMR spectrum 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(8''-quinolyl)-2'-dG (3d) in DMSO-d$_6$.

Figure B - 34 $^{13}$C NMR spectrum 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(8''-quinolyl)-2'-dG (3d) in DMSO-d$_6$.

Figure B - 35 $^1$H NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (3e) in CDCl$_3$.

Figure B - 36 $^{13}$C NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-Benzothienyl)-2'-dG (3e) in CDCl$_3$.

Figure B - 37 $^1$H NMR spectrum 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (3f) in CDCl$_3$.

Figure B - 38 $^{13}$C NMR spectrum 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (3f) in CDCl$_3$.

Figure B - 39 $^1$H NMR spectrum DMPx-OMe in CDCl$_3$.

Figure B - 40 $^{13}$C NMR spectrum DMPx-OMe in CDCl$_3$.

Figure B - 41 $^1$H NMR spectrum DMPx-Cl in CD$_2$Cl$_2$.

Figure B - 42 $^{13}$C NMR spectrum DMPx-Cl in CD$_2$Cl$_2$.

Figure B - 43 $^1$H NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(2''-furanlyl)-2'-dG (5a) in CD$_2$Cl$_2$.

Figure B - 44 $^{13}$C NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(2''-furanlyl)-2'-dG (5a) in CD$_2$Cl$_2$.

Figure B - 45 $^1$H NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (5b) in CD$_2$Cl$_2$.

Figure B - 46 $^1$H NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (5b) in CD$_2$Cl$_2$.

Figure B - 47 $^1$H NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-
2′-dG (5e) in DMSO-d$_6$.

Figure B-48 $^{13}$C NMR spectrum of 5′-O-DMPx-N$_2$-(dimethylformamidyl)-8-(4′-cyanophenyl)-2′-dG (5e) in DMSO-d$_6$.

Figure B-49 $^1$H NMR spectrum 5′-O-DMPx-N$_2$-(dimethylformamidyl)-dG (5g) in CDCl$_3$.

Figure B-50 $^{13}$C NMR spectrum 5′-O-DMPx-N$_2$-(dimethylformamidyl)-dG (5g) in CDCl$_3$.

Figure B-51 $^1$H NMR spectrum 5′-O-DMPx-N$_2$-(acetyl)-2′-dC (5h) in CD$_2$Cl$_2$.

Figure B-52 $^{13}$C NMR spectrum of 5′-O-DMPx-N$_2$-(acetyl)-2′-dC (5h) in CD$_2$Cl$_2$.

Figure B-53 $^1$H NMR spectrum 5′-O-DMPx-2′-dT (5i) in DMSO-d$_6$.

Figure B-54 $^{13}$C NMR spectrum 5′-O-DMPx-2′-dT (5i) in DMSO-d$_6$.

Figure B-55 $^1$H NMR spectrum 5′-O-DMPx-N$_2$-(dimethylformamidyl)-dG (5j) in CDCl$_3$.

Figure B-56 $^{13}$C NMR spectrum 5′-O-DMPx-N$_2$-(dimethylformamidyl)-dG (5j) in CDCl$_3$.

Figure B-57 $^1$H NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (4a) in CDCl$_3$.

Figure B-58 $^{13}$C NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (4a) in CDCl$_3$.

Figure B-59 $^{31}$P NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (4a) in CDCl$_3$.

Figure B-60 $^1$H NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (6a) in CD$_2$Cl$_2$.

Figure B-61 $^{13}$C NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (6a) in CD$_2$Cl$_2$.

Figure B-62 $^{31}$P NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$_2$-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (6a) in CD$_2$Cl$_2$.

Figure B-63 $^1$H NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-phenyl-2′-dG (4b) in CDCl$_3$.

Figure B-64 $^{13}$C NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-phenyl-2′-dG (4b) in CDCl$_3$.

Figure B-65 $^{31}$P NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N$_2$-(dimethylformamidyl)-8-phenyl-2′-dG (5b) in CDCl$_3$.

Figure B-66 $^1$H NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$_2$-(dimethylformamidyl)-8-phenyl-2′-dG (6b) in CD$_2$Cl$_2$.

Figure B-67 $^{13}$C NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-
DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (6b) in CD$_2$Cl$_2$.

Figure B - 68 $^3$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (6b) in CD$_2$Cl$_2$.

Figure B - 69 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (4c) in CDCl$_3$.

Figure B - 70 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (4c) in CDCl$_3$.

Figure B - 71 $^3$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (6c) in CD$_2$OD.

Figure B - 72 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (6c) in CD$_2$OD.

Figure B - 73 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (6c) in CD$_2$OD.

Figure B - 74 $^3$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4'-cyanophenyl)-2'-dG (6c) in CD$_2$OD.

Figure B - 75 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT $N^2$-(dimethylformamidyl)-8-(8'-quinolyl)-2'-dG (4d) in DMSO-d$_6$.

Figure B - 76 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT $N^2$-(dimethylformamidyl)-8-(8'-quinolyl)-2'-dG (4d) in CDCl$_3$.

Figure B - 77 $^3$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(8'-quinolyl)-2'-dG (4d) in DMSO-d$_6$.

Figure B - 78 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2'-benzothienyl)-2'-dG (4e) in Acetone-d$_6$.

Figure B - 79 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2'-benzothienyl)-2'-dG (4e) in Acetone-d$_6$.

Figure B - 80 $^3$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2'-benzothienyl)-2'-dG (4e) in Acetone-d$_6$.

Figure B - 81 $^1$H NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (4f) in CD$_2$Cl$_2$.

Figure B - 82 $^{13}$C NMR of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (4f) in CD$_2$Cl$_2$.

Figure B - 83 $^3$P NMR of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (4f) in CD$_2$Cl$_2$.

Figure B - 84 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-2'-dG in (6g) CD$_2$Cl$_2$. xxii
Figure B - 85 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dG (6g) in CD$_2$Cl$_2$.

Figure B - 86 $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dG (6g) in CD$_2$Cl$_2$.

Figure B - 87 $^1$H NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^4$-(acetyl)-2'-dC (6h) in CDCl$_3$.

Figure B - 88 $^{13}$C NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^4$-(acetyl)-2'-dC (6h) in CDCl$_3$.

Figure B - 89 $^{31}$P NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^4$-(acetyl)-2'-dC (6h) in CDCl$_3$.

Figure B - 90 $^1$H NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-2'-dT (6i) in CD$_2$Cl$_2$.

Figure B - 91 $^{13}$C NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-2'-dT (6i) in CD$_2$Cl$_2$.

Figure B - 92 $^{31}$P NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-2'-dT (6i) in CD$_2$Cl$_2$.

Figure B - 93 $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$.

Figure B - 94 $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$.

Figure B - 95 $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$.

Figure C - 1 Mass spectrum of 12mer NarI Fur dG, obtained with an ESI source operated in negative mode.

Figure C - 2 Mass spectrum of 12mer NarI Bth dG, obtained with an ESI source operated in negative mode.

Figure C - 3 Mass spectrum of 12mer NarI CNPh dG, obtained with an ESI source operated in negative mode.

Figure C - 4 Mass spectrum of 12mer NarI QdG, obtained with an ESI source operated in negative mode.

Figure C - 5 Mass spectrum of 12mer NarI Py dG, obtained with an ESI source operated in negative mode.

Figure C - 6 Mass spectrum of 12mer NarI Ph dG, obtained with an ESI source operated in negative mode.
Figure C - 7  Mass spectrum of 22-mer \textit{NarI} \textsubscript{Fur}dG, obtained with an ESI source operated in negative mode.

Figure C - 8  Mass spectrum of 22-mer \textit{NarI} \textsubscript{Bth}dG, obtained with an ESI source operated in negative mode.

Figure C - 9  Mass spectrum of 22-mer \textit{NarI} \textsubscript{Q}dG, obtained with an ESI source operated in negative mode.

Figure C - 10 Mass spectrum of 22-mer \textit{NarI} \textsubscript{Py}dG, obtained with an ESI source operated in negative mode.

Figure C - 11 Mass spectrum of 22-mer \textit{NarI} \textsubscript{Ph}dG, obtained with an ESI source operated in negative mode.

Figure C - 12 Mass spectrum of TBA \textsubscript{Fur}dG, obtained with an ESI source operated in negative mode.

Figure C - 13 Mass spectrum of TBA \textsubscript{Bth}dG, obtained with an ESI source operated in negative mode.

Figure C - 14 Mass spectrum of TBA \textsubscript{CNPh}dG, obtained with an ESI source operated in negative mode.

Figure C - 15 Mass spectrum of TBA \textsubscript{Q}dG, obtained with an ESI source operated in negative mode.

Figure C - 16 Mass spectrum of TBA \textsubscript{Py}dG, obtained with an ESI source operated in negative mode.

Figure C - 17 Mass spectrum of TBA \textsubscript{Fur}dG + \textsubscript{Py}dG, obtained with an ESI source operated in negative mode.

Figure C - 18 Mass spectrum of TBA \textsubscript{Fur}dG + \textsubscript{CNPh}dG, obtained with an ESI source operated in negative mode.

Figure C - 19 Mass spectrum of HTelo22 \textsubscript{Fur}dG obtained with an ESI source operated in negative mode.
<table>
<thead>
<tr>
<th>Scheme 1 - 1</th>
<th>Slippage model initially proposed by Streisinger and coworkers for the <em>NarI</em> Recognition sequence.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme 2 - 1</td>
<td>Deglycosylation (depurination) of a dG nucleobase within DNA and subsequent base catalyzed strand scission.</td>
</tr>
<tr>
<td>Scheme 2 - 2</td>
<td>Mechanism of the acid catalyzed deglycosylation of dG.</td>
</tr>
<tr>
<td>Scheme 2 - 3</td>
<td>Catalytic cycle of the Suzuki-Miyuara cross-coupling reaction.</td>
</tr>
<tr>
<td>Scheme 2 - 4</td>
<td>Solid-phase four step phosphoramidite oligodeoxynucleotide synthesis cycle.</td>
</tr>
<tr>
<td>Scheme 2 - 5</td>
<td>Outline of the synthesis of various 8-aryl-dG nucleoside phosphoramidites.</td>
</tr>
<tr>
<td>Scheme 2 - 6</td>
<td>Synthesis of DMPx – Cl</td>
</tr>
<tr>
<td>Scheme 3 - 1</td>
<td>Thermal melting analysis of 8-aryl-dG adducts within various <em>NarI</em> duplexes. (a) <em>NarI</em>(X):<em>NarI</em>(C); (b) <em>NarI</em>(X):<em>NarI</em>(G) (c) <em>NarI</em>(X):<em>NarI</em>(10mer) <em>NarI</em>(X):<em>NarI</em>(THF). Dotted black line X = dG, Red line X = <em>Fur</em>dG, Green line X = <em>Bth</em>dG, Blue line X = <em>Q</em>dG, Orange Line X = <em>Py</em>dG, Purple line X = <em>CNPh</em>dG, solid black line X = <em>Ph</em>dG.</td>
</tr>
<tr>
<td>Scheme 4 - 1</td>
<td>Visualization of the SELEX process.</td>
</tr>
<tr>
<td>Scheme 5 - 1</td>
<td>The end replication problem. DNA at the end of telomeres exists as a single-stranded 3′ overhang. Parental duplex (black) is unwound at the replication fork and replication proceeds continuously on the leading strand and discontinuously on the lagging strand. RNA primers are shown in red while the generated DNA fragments are shown in blue.</td>
</tr>
<tr>
<td>Scheme 5 - 2</td>
<td>Quadruplex structure of the various HTelo22 topologies depicting position of modification by <em>Fur</em>dG.</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-AP</td>
<td>2-aminopurine</td>
</tr>
<tr>
<td>8-Br-dG</td>
<td>8-bromo-2'-deoxyguanosine</td>
</tr>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>A</td>
<td>adenine</td>
</tr>
<tr>
<td>A.U.</td>
<td>arbitrary unit</td>
</tr>
<tr>
<td>AAF</td>
<td>2-(acetylamino)fluorene</td>
</tr>
<tr>
<td>AAF-dG</td>
<td>8-(2-acetyaminofluorene)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>ACN</td>
<td>acetonitrile</td>
</tr>
<tr>
<td>AF</td>
<td>aminofluorene</td>
</tr>
<tr>
<td>AFB1</td>
<td>aflatoxin B1</td>
</tr>
<tr>
<td>AF-dG</td>
<td>8-aminofluorene-2'-deoxyguanosine</td>
</tr>
<tr>
<td>AN-dG</td>
<td>8-aminophenyl-2'-deoxyguanosine</td>
</tr>
<tr>
<td>AP-dG</td>
<td>8-(aminopyrene)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>AQ</td>
<td>antiquinoid</td>
</tr>
<tr>
<td>aq.</td>
<td>aqueous</td>
</tr>
<tr>
<td>ARCO</td>
<td>aryloxycarbonyl</td>
</tr>
<tr>
<td>B-type</td>
<td>major groove binding conformation - B-conformation</td>
</tr>
<tr>
<td>BER</td>
<td>base excision repair</td>
</tr>
<tr>
<td>BP</td>
<td>benzo[a]pyrene</td>
</tr>
<tr>
<td>bs</td>
<td>broad singlet</td>
</tr>
<tr>
<td>Bth-dG</td>
<td>8-(2''-benzothienyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>calcd</td>
<td>calculated</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
</tbody>
</table>
CT charge transfer
Cy cyanine
d doublet
D-A donor-acceptor
dA 2'-deoxyadenosine
dC 2'-deoxycytidine
DCA dichloroacetic acid
DCM dichloromethane
DFT density functional theory
dG 2'-deoxyguanosine
DMABN dimethylaminobenzonitrile
DMF dimethylformamide
DMPx 2,7-dimethylpixyl
DMSO dimethylsulfoxide
DMT 4,4'-dimethoxytrityl
DMT-Cl 4,4'-dimethoxytrityl chloride
D_N leaving group dissociation
DNA deoxyribonucleic acid
dR 2'-deoxyribose
dsDNA double strand DNA (duplex DNA)
dT 2'deoxythymidine
_E.Coli_ *Escherichia Coli_ 
E_{p/2} half-peak oxidation potential
EPA Environmental Protection Agency
eq. equivalent(s)
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESI-MS</td>
<td>electrospray ionization mass spectrometry</td>
</tr>
<tr>
<td>Et&lt;sub&gt;3&lt;/sub&gt;N</td>
<td>triethylamine</td>
</tr>
<tr>
<td>FAAF-dG</td>
<td>8-(2-acetyl-4-fluoroaminofluorene)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>FAF-dG</td>
<td>8-(N-2-fluoroaminofluorene)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>Fur&lt;sup&gt;d&lt;/sup&gt;G</td>
<td>8-(2''-furyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>GBA</td>
<td>glycosidic bond angle</td>
</tr>
<tr>
<td>HOMO</td>
<td>highest occupied molecular orbital</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>high resolution mass spectrometry</td>
</tr>
<tr>
<td>HRP</td>
<td>horse radish peroxidase</td>
</tr>
<tr>
<td>HTelo</td>
<td>Human telomeric sequence 5′-AGGG(TTAGGG)&lt;sub&gt;3&lt;/sub&gt;-3′</td>
</tr>
<tr>
<td>IARC</td>
<td>International Agency for Research on Cancer</td>
</tr>
<tr>
<td>ICD</td>
<td>induced circular dichroism</td>
</tr>
<tr>
<td>ICT</td>
<td>internal charge transfer</td>
</tr>
<tr>
<td>I&lt;sub&gt;rel&lt;/sub&gt;</td>
<td>relative emission intensity</td>
</tr>
<tr>
<td>IQ</td>
<td>2-amino-3-methyl-imidazo[4,5-f]quinoline</td>
</tr>
<tr>
<td>IUPAC</td>
<td>International Union of Pure and Applied Chemistry</td>
</tr>
<tr>
<td>J</td>
<td>Joules</td>
</tr>
<tr>
<td>K</td>
<td>Kelvin</td>
</tr>
<tr>
<td>k&lt;sub&gt;1&lt;/sub&gt;</td>
<td>rate constant for the glycosidic bond cleavage of the monoprotonated species</td>
</tr>
<tr>
<td>k&lt;sub&gt;2&lt;/sub&gt;</td>
<td>rate constant for the glycosidic bond cleavage of the monoprotonated species</td>
</tr>
<tr>
<td>K&lt;sub&gt;a&lt;/sub&gt;</td>
<td>acid dissociation constant or association equilibrium constant</td>
</tr>
<tr>
<td>k&lt;sub&gt;obs&lt;/sub&gt;</td>
<td>observed first-order rate constant</td>
</tr>
<tr>
<td>l</td>
<td>path length</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LE</td>
<td>locally excited</td>
</tr>
<tr>
<td>LUMO</td>
<td>lowest unoccupied molecular orbital</td>
</tr>
<tr>
<td>m</td>
<td>multiplet</td>
</tr>
<tr>
<td>MD</td>
<td>molecular dynamics</td>
</tr>
<tr>
<td>mer</td>
<td>denotes length of an oligonucleotide</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>NER</td>
<td>nucleotide excision repair</td>
</tr>
<tr>
<td>N.H.E.</td>
<td>normal hydrogen electrode</td>
</tr>
<tr>
<td>NarI</td>
<td>5'-CTCGGC GCC CATC-3'</td>
</tr>
<tr>
<td>NarI(10)</td>
<td>5'-GATGG CCGAG-3'</td>
</tr>
<tr>
<td>NarI(C)</td>
<td>5'-GATGG CCGAG-3'</td>
</tr>
<tr>
<td>NarI(X)</td>
<td>5'-CTCGGC XCCATC-3'</td>
</tr>
<tr>
<td>NarI(G)</td>
<td>5'-GATGG GCCGAG-3'</td>
</tr>
<tr>
<td>NarI(THF)</td>
<td>5'-GATGG-THF-GCCGAG-3'</td>
</tr>
<tr>
<td>NBS</td>
<td>N-bromosuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>OAc</td>
<td>acetate</td>
</tr>
<tr>
<td>OTA</td>
<td>ochratoxin A</td>
</tr>
<tr>
<td>PAH</td>
<td>polyaromatic hydrocarbon</td>
</tr>
<tr>
<td>PCP</td>
<td>pentachlorophenol</td>
</tr>
<tr>
<td>PET</td>
<td>photoinduced electron transfer</td>
</tr>
<tr>
<td>PhdG</td>
<td>8-(phenyl)- 2'-deoxyguanosine</td>
</tr>
<tr>
<td>pK_a</td>
<td>ionization constant</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Px</td>
<td>9-phenylxanthen-9-yl</td>
</tr>
<tr>
<td>PydG</td>
<td>8-(pyrenyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>Q</td>
<td>quinoid or quencher</td>
</tr>
<tr>
<td>QdG</td>
<td>8-(8''-quinolyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>RBF</td>
<td>round bottomed flask</td>
</tr>
<tr>
<td>RET</td>
<td>resonance energy transfer</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>s</td>
<td>seconds or singlet</td>
</tr>
<tr>
<td>S</td>
<td>base displaced stacked conformation - stacked</td>
</tr>
<tr>
<td>So</td>
<td>ground state</td>
</tr>
<tr>
<td>Sn, N&gt;0</td>
<td>excited state</td>
</tr>
<tr>
<td>SNAr</td>
<td>nucleophilic aromatic substitution</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single stranded DNA</td>
</tr>
<tr>
<td>S-type</td>
<td>stacked conformation</td>
</tr>
<tr>
<td>T</td>
<td>thymine or temperature</td>
</tr>
<tr>
<td>t1/2</td>
<td>half-life</td>
</tr>
<tr>
<td>TBA</td>
<td>thrombin binding aptamer 5'-GGTTGGTGTGGTTGG-3'</td>
</tr>
<tr>
<td>TBAF</td>
<td>tetrabutylammonium fluoride</td>
</tr>
<tr>
<td>TEA</td>
<td>triethylamine</td>
</tr>
<tr>
<td>TEAA</td>
<td>triethylammonium acetate</td>
</tr>
<tr>
<td>TEBP</td>
<td>telomere end binding proteins</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TICT</td>
<td>twisted intramolecular charge transfer</td>
</tr>
<tr>
<td>Tm</td>
<td>thermal melting</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>TPPTS</td>
<td>tris(3-sulfophenyl)phosphine trisodium salt hydrate</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>UV-vis</td>
<td>ultraviolet-visible</td>
</tr>
<tr>
<td>V</td>
<td>voltage</td>
</tr>
<tr>
<td>W-C</td>
<td>Watson-Crick</td>
</tr>
<tr>
<td>W-type</td>
<td>minor groove binding conformation – wedge conformation</td>
</tr>
<tr>
<td>ΔG</td>
<td>Gibb's free energy of activation</td>
</tr>
<tr>
<td>ΔH</td>
<td>enthalpy of activation</td>
</tr>
<tr>
<td>ΔS</td>
<td>entropy of activation</td>
</tr>
<tr>
<td>Δν</td>
<td>Stokes shift</td>
</tr>
<tr>
<td>δ</td>
<td>NMR chemical shift</td>
</tr>
<tr>
<td>ε</td>
<td>dielectric constant or molar extinction coefficient</td>
</tr>
<tr>
<td>Θ</td>
<td>molar ellipticity or dihedral twist angle (∠(C\text{11}C\text{10}C\text{8}N\text{9}))</td>
</tr>
<tr>
<td>λ</td>
<td>wavelength</td>
</tr>
<tr>
<td>μs</td>
<td>dipole moment</td>
</tr>
<tr>
<td>Φf</td>
<td>fluorescence quantum yield</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction
1.1. Background Information

1.1.1. Historical Importance

DNA is considered to be one of the most versatile macromolecules within a cell. It is involved in reactions essential to the propagation and maintenance of life, and is responsible for the storage and transfer of genetic information. DNA was first discovered in 1869 by Friedrich Miescher who termed it “nuclein;” its function, however, remained unknown.\(^1\) Within the next 70 years, major efforts aimed at determining the primary structure of DNA; its role as genetic material was not realized until 1944 when Oswald Avery discovered its role in inheritance within bacteria.\(^2\) However, information on the primary structure alone could not explain how DNA functioned as a genetic material. It became apparent that information regarding the tertiary structure of DNA was needed to answer these questions. It took until 1953 for the first correct three dimensional-model of DNA to be proposed by James Watson and Francis Crick, based on extensive work by Rosalind Franklin and Maurice Wilkins.\(^1\) This era is often considered to be the beginning of modern molecular biology.

Fast forward 60 years and the understanding of the structural variations and biological functions of DNA have grown exponentially. Technological advances have allowed DNA sequencing projects to be completed that have mapped the entire human genome, as well as those of numerous other species.\(^3\) Every day researchers are getting a better grasp on how DNA complexes with various proteins, drugs, metals and other molecules. These interactions have deepened our understanding of DNA damage, the resulting processing/repair of such lesions, and has leant justifications to mutations and the onset of cancer.

More recently, DNA’s role as a functioning nanoscaffold, for a wide range of applications in genetic engineering and nanotechnology, is starting to be realized. The term “aptamer” was termed for a single structured nucleic acid capable of binding to a specific target. The first aptamer,
although it was not referred to as such, was created just over 40 years ago in the late 1960s by Sol
Spiegelman.\textsuperscript{4} Aptamer science has now reached maturity; it is having a growing impact on biology
and medicine. In December 2004, the first aptamer compound was approved for clinical use;
Macugen has become a preferred treatment for the neovascular form of age-related macular
degeneration.\textsuperscript{5} Aptamers are being developed for various therapeutic applications, medical imaging,
clinical diagnostics, drug target validation, biosensor applications, and process chemistry.\textsuperscript{6}

It will be interesting to see how our understanding of DNA and its applications evolve over
the next 10 years.

1.1.2. Primary/Secondary Structure of Nucleic Acids

Deoxyribonucleic acid (DNA) is a biopolymer essential to life. Structurally, it is comprised
of four monomeric units, termed deoxyribonucleotide units (Figure 1 - 1). These
deoxyribonucleotide units consist of (1) a nucleobase (2) a deoxyribose sugar and (3) a
phosphodiester linkage; when the phosphates are absent from a nucleotide the unit is termed a
nucleoside.\textsuperscript{7} The nucleobases that constitute nucleosides are aromatic, planar, heterocyclic
molecules derived from either a purine or a pyrimidine. The most notable nucleobases classified as
purines are guanine (G) and adenine (A), while those classified as pyrimidines are cytosine (C) and
thymine (T). The connectivity of the nucleobase to the 2’-deoxyribose sugar is through a \( \beta \)-N-
glycosidic bond; this glycosidic bond results in puckering of the sugar. This connectivity results in a
DNA nucleoside, for which the monomers are termed deoxyguanosine, deoxyadenosine, thymidine,
and deoxycytidine. The numbering conventions for the pentose ring and nucleobase component are
depicted below in Figure 1 - 1.
1.1.3. Properties of DNA

1.1.3.1. Ionisation and Tautomerism

Nucleosides are relatively hydrophobic and insoluble in water (soluble in water to approximately 1% w/v varying with size). Within DNA, they are rendered soluble by the connection of the 3′-OH and 5′-OH positions of the sugars to the phosphate backbone, which is negatively charged at physiological pH. Important physical characteristics of nucleosides and nucleotides are their charge, tautomeric structure and ability to donate and/or accept hydrogen bonds.

As can be seen in Table 1 - 1 all of the nucleobases themselves are uncharged in the physiological range. This is also the case for the pentose sugars, the ribose of 2′- and 3′- OH diols has a pKa of approximately 12, while isolated hydroxyl groups have pKa’s around 15. The proximity of the negatively charged phosphate residues has a secondary effect on the pKa values of the nucleobases, making the ring nitrogens more basic (ΔpKa ≈ +0.4) and the amine protons less acidic (ΔpKa ≈ +0.6).
Table 1 - 1: The pKa values for bases as nucleosides and nucleotides.

<table>
<thead>
<tr>
<th>Base</th>
<th>Nucleoside pKa</th>
<th>3'-Nucleotide</th>
<th>5'-Nucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine (N-1)</td>
<td>3.63</td>
<td>3.74</td>
<td>3.74</td>
</tr>
<tr>
<td>Cytosine (N-3)</td>
<td>4.11</td>
<td>4.30</td>
<td>4.56</td>
</tr>
<tr>
<td>Guanine (N-7)</td>
<td>2.20</td>
<td>2.30</td>
<td>2.40</td>
</tr>
<tr>
<td>Guanine (N-1)</td>
<td>9.5</td>
<td>9.36</td>
<td>9.40</td>
</tr>
<tr>
<td>Thymine(N-3)</td>
<td>9.8</td>
<td>--</td>
<td>10.00</td>
</tr>
<tr>
<td>Uracil (N-3)</td>
<td>9.25</td>
<td>9.43</td>
<td>9.50</td>
</tr>
</tbody>
</table>

It is important to note that protonation preferentially occurs on one of the ring nitrogens rather than the exocyclic amino groups for the three amino bases (A, C and G). This is because protonation of the ring nitrogens does not interfere with delocalisation of the NH₂ electron lone pair into the aromatic system. The C-NH₂ bonds of A (N̄^6), C(N̄^4) and G (N̄^2) have about 40-50% double bond order and are about 1.34 Å in length.⁸

1.1.3.2. Hydrogen Bonding

NMR and IR spectroscopies show that the major bases exist overwhelmingly in the amino- and keto tautomeric forms at physiological pH.⁸ This tautomeric equilibrium results in the nucleobases within duplex DNA establishing a network of distinct hydrogen bonding interactions with one another. These hydrogen bonding interactions allow the variation of linearly connected nucleotide units to generate a sequence that permits DNA to encode the genetic instructions used in the development and functioning of all known living organisms. DNA is also said to lie at the heart of the central dogma of molecular biology, as it provides the blueprint that can be transcribed into RNA and further translated into protein’s by various components of molecular machinery.⁹

The H-bonding scenario is depicted in Figure 1 - 2 where A bonds to T via two H-bonds and G bonds to C via three H-bonds. The atoms that are implicated in Watson-Crick (W-C) H-bonding comprise what is commonly referred to as the W-C face of the respective base. As can be seen,
these specific H-bonding interactions are what account for the specificity and discrimination of one nucleoside for its partner. In planar base pairs, the hydrogen bonds join nitrogen and oxygen atoms that are 2.8–2.95 Å apart. This geometry gives a C-1’ - C-1’ distance of 10.60 ± 0.15 Å with an angle of 68 ± 2° between the two glycosylic bonds for both the A-T and the C-G base pairs. As a result of this isomorphous geometry, the four base pair combinations A-T, T-A, C-G and G-C can all be built into the same regular framework of the DNA duplex.  

![Figure 1 - 2: H-bonding patterns of the Watson-Crick Base Pairs](image)

Another H-bonding scenario that exists besides the standard W-C type H-bonding is Hoogsteen H-bonding. Hoogsteen H-bonding results in base pairing defined by the participation of the N7-, and O6- or N6- atoms of a purine, in H-bonding with the W-C face of the complementary base, which can be a purine or pyrimidine (Figure 1 - 3). The Hoogsteen base-pair of G with C is most stable at low pH, as a protonated cytosine is required for H-bonding. In the base-pair, the pKa of cytosine is > 7.5, which is considerably higher than its normal value of 4.11.

![Figure 1 - 3: H-bonding patterns of Hoogsteen base pairing](image)
The linear connectivity of successive nucleotides to form DNA occurs through covalently linking the 5' phosphate group of one nucleotide with the 3' hydroxyl group of the next nucleotide (Figure 1 - 4). This bond between the C3' on one pentose sugar with the C5' on the adjoining pentose is termed a phosphodiester bond. This connectivity creates an oligonucleotide, which is defined as the polymer of DNA containing fewer than 100 monomer building blocks; sequences that are longer are termed polynucleotides. The term ‘mer’ is commonly employed to describe the length of an oligonucleotide, referring to the number of nucleotides it is comprised of (i.e. an 11mer oligonucleotide would contain 11 nucleotides). The tetradeoxynucleotide shown in Figure 1 - 4 is adenylyl-3',5'-thymidinyl-3',5'-guanylyl-3',5'-cytidine and is abbreviated d(ApTpGpC) or d(ATGC). By convention, an oligonucleotide sequence is written from the 5' end to the 3' end.

![Figure 1 - 4: The covalent backbone of DNA, showing the linkage of successive nucleotide units by phosphodiester bonds.](image)

1.1.3.3. Sugar-Phosphate Backbone and the Glycosyl Bond

There are seven torsion angles that describe the secondary structure of a nucleotide, termed \( \alpha, \beta, \gamma, \delta, \epsilon, \zeta \) and \( \chi \) (Figure 1 - 5). Most torsion angles describe the orientation of the sugar phosphate backbone, while the last torsion angle \( \chi \) refers to the rotation around the glycosidic bond.
Because many of these torsional angles are interdependent, one can simply describe the shapes of the nucleotides in terms of four parameters: the conformation of the glycosidic bond, the sugar pucker, the orientation of C4′ and C5′ and the shape of the phosphate bonds. Here, the focus will be on the first two categories.

![Torsional angles for a nucleotide within the DNA duplex.](image)

Steric interactions greatly influence the rotation around the glycosyl bond, especially for the purine bases. The most energetically permissible orientations for the glycosyl bond are syn and anti. Anti conformations are defined by the dihedral angle $\chi = 180 \pm 90^\circ$ ([O4′C1′N9C4 for purines and [O4′C1′N1C2] for pyrimidines) while for syn conformations $\chi = 0 \pm 90^\circ$. One can see that in the anti conformation for purine bases the bulk of the nitrogenous base is pointed away from the sugar, which minimizes steric clash (Figure 1-6). The amount of steric hindrance that exists between the nitrogenous base and the sugar is also affected by the conformation of the sugar, i.e. the sugar pucker. Sugar pucker describes the flexibility of the non-planar deoxyribose ring. The most common puckers available for DNA are the C2′-endo (typically found in anti structures) and the C3′-endo conformations with the C2′-endo conformation being preferred (Figure 1-6) for β-DNA. These conformations are in rapid equilibrium in solution, and are separated by an energy barrier of less than 20 kJ/mol.
The syn/anti equilibrium becomes more complicated within the duplex, resulting in three major conformations which exist in equilibrium,\textsuperscript{11,12} this is a product of the complexity of the duplex environment. Factors affecting conformations include: the stability of the W-C H-bonding interaction vs. possible Hoogsteen H-bonding,\textsuperscript{13} as well as the π stacking ability,\textsuperscript{14} which is a factor of its size and planarity.\textsuperscript{15}

1.1.3.4. Oxidative Properties

The oxidative properties of the nucleosides are well established. Of the four nucleosides, dG has the highest electron density and owns the lowest oxidation potential; therefore, most oxidizing agents attack guanine with a greater frequency than the other nucleobases. Guanine has an oxidation potential of 1.29 V versus normal hydrogen electrode (NHE), while adenine, cytosine and thymine have oxidation potentials of 1.42, 1.6 and 1.7 V/NHE, respectively.\textsuperscript{16}

1.1.3.5. Photophysical Properties

It is evident from the structure of the nucleobases that they are highly conjugated molecules; this high degree of conjugation gives rise to multiple resonance structures resulting in most of the bonds owning partial double-bond character. Nucleobases are chromophores and absorb ultraviolet light through excitation of electrons in a ground state π orbital to an excited state π∗ orbital. Neither
the pentose nor the phosphate components of nucleotides show any significant UV absorption above 230 nm. Therefore, both nucleosides and nucleotides have UV absorption profiles similar to those of their constituent nucleobases and absorb strongly with $\lambda_{\text{max}}$ values close to 260 nm and molar extinction coefficients of around $10^4$. Table 1-2 depicts the specific absorption wavelength maxima as well as molar extinction coefficients ($\varepsilon$), which describe how strongly a molecule absorbs light at a given wavelength.

Table 1-2: Photophysical parameters of nucleosides.

<table>
<thead>
<tr>
<th>Adduct</th>
<th>$\lambda_{\text{abs/max}}$ (nm)$^a$</th>
<th>$\varepsilon_{\text{max}}$ (x10$^3$ M$^{-1}$cm$^{-1}$)$^a$</th>
<th>$\lambda_{\text{em}}$ (nm)</th>
<th>$\Phi_{\text{f}}$ (x10$^{-4}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dG</td>
<td>253</td>
<td>13.8</td>
<td>334</td>
<td>0.97</td>
</tr>
<tr>
<td>dA</td>
<td>269</td>
<td>14.9</td>
<td>307</td>
<td>0.86</td>
</tr>
<tr>
<td>dT</td>
<td>267</td>
<td>9.86</td>
<td>330</td>
<td>1.32</td>
</tr>
<tr>
<td>dC</td>
<td>271</td>
<td>9.30</td>
<td>328</td>
<td>0.89</td>
</tr>
</tbody>
</table>

$^a$ Data recorded in aqueous solutions, as taken from Reference.$^{17}$

The absorbance values given above correspond to isolated nucleoside bases measured in dilute aqueous solution. When the nucleosides are in close proximity to neighbouring bases, as is the case in ordered secondary structures of oligo- and poly-nucleotides, there is a reduction in the intensity of UV absorption by up to 30%, which is known as hypochromicity. In such ordered structures, the bases can stack face-to-face and thus share π-π electron interactions that profoundly affect the transition dipoles of the bases.$^8$ This phenomenon is reversed on unstacking of the bases and is commonly utilized for photophysical characterization of oligonucleotides via thermal melting ($T_m$) and/or circular dichroism (CD).

Table 1-2 shows that, although the nucleoside derivatives have fairly large extinction coefficients, their fluorescent quantum yields are very low. A molecule’s quantum yield gives a measure of the efficiency of the fluorescence process. It is defined as the ratio of the number of photons emitted to the number of photons absorbed, and is obtained by comparing the fluorescence intensity of the molecule with that of a reference molecule with a known quantum yield.$^{18}$
Schematically, Jablonski diagrams are utilized to illustrate the overall photochemistry; an example of a Jablonski diagram is depicted in Figure 1 - 7.

In these diagrams transitions between states are depicted as vertical lines. Most absorption transitions occur in approximately $10^{15}$ s, a time too short for significant displacement of nuclei (Frank-Condon principle). Fluorescence emission processes take place on a slower time scale (approximately $10^9$ s) which allows a much wider range of interactions to influence the fluorescence emission spectrum. Absorption occurs mostly from molecules in the ground state with differing vibrational energy levels to an excited electronic state (either $S_1$ or $S_2$) with some higher vibrational level. Most molecules rapidly relax to the lowest vibrational level of an electronic state through internal conversion in $10^{12}$ s. Due to excited state fluorescence lifetimes typically having values around $10^9$ s, internal conversion is complete before fluorescence occurs.

In comparing the techniques used for studying nucleic acid-containing systems (e.g. NMR, X-ray crystallography, electrophoresis, foot-printing, calorimetry, enzymatic methods and UV-Vis absorption), fluorescence is one of the most sensitive, versatile and easily accessible. It is used commonly in single-molecule real-time dynamics of nucleic acids as well as proteins, cell...
microscopy, nucleic acid detection and nucleic acid-protein interaction measurements. Thus, fluorescence studies at both the ensemble and single-molecule level provide researchers with biophysical data on nucleic acids, and will increase in importance in the expanding field of DNA nanotechnology as these constructs keep decreasing in dimensions.

The low fluorescence quantum yield of the nucleobases is attributed to the excited-state lifetimes in the order of picoseconds. Therefore, quenching of the excited state occurs by internal conversion to the ground state (which is on the same order of magnitude \(10^{-12}\) s). This short excited state lifetime/internal conversion relaxation is considered vital to the preservation of life on earth as the short lifetime protects the nucleosides from photochemistry.

As a result of the poor fluorescent quantum yields of the nucleosides significant efforts have been made to develop fluorescent nucleic acid base analogues suitable for studying nucleic acid-containing systems. Typical strategies involve introducing a fluorophore either covalently or non-covalently into the system. The covalently bound fluorophore analogues are considered as either external modifications or internal modifications.

External modifications refer to modifying nucleic acids by covalently attaching fluorophores to the backbone at the end of, or within an oligonucleotide sequence, but outside the actual base stack. This is currently the most common way of labelling DNA and is used with commercially available fluorescein, rhodamine derivatives and cyanine dyes (Cy) (Figure 1 - 8). These fluorophores are extremely bright and, thus, especially useful as probe molecules in gel electrophoresis experiments, different fluorescence microscopy techniques, or single- or few-molecule experiments. The molar absorptivity of a dye frequently changes after attachment to biomolecules, or changes when dissolved in solutions of different solvents salt concentrations, or with varying pH.
In an internal modification, a covalently attached fluorophore replaces the nucleobase inside the base stack. Typically, planar aromatic compounds are attached at the nucleobase position of the deoxyribose moiety. These planar aromatics are generally different in shape and size and lack the possibility to form hydrogen bonds with the base on the opposite strand. The most widely used nucleobase analogue is 2-aminopurine (2-AP) (Figure 1 - 9 (a)); 2-AP has been used extensively in probing structure and dynamics of DNA.

Pteridines developed by Hawkins and Pfeiderer are also interesting and frequently used as hybridization probes to monitor enzyme-catalyzed reactions of DNA. The most promising pteridine analogues include the guanine analogues 3-MI and 6MI (Figure 1 - 9 (b)) and the adenine analogues 6-MAP and DMAP (Figure 1 - 9 (c)). Kool et al. have replaced nucleobases with polycyclic hydrocarbons such as pyrene, phenanthrene and stilbene which, due to their strong stacking interactions, are incorporated with minor perturbations to the

Figure 1 - 8: Various external fluorescence modifications linked to the terminus of an oligonucleotide. (a) Fluoroscein 6-FAM; (b) NHS-Rhodamine; (c) Cy3 dye.
DNA structure, even though H-bonding is not possible to their complement.\textsuperscript{34,35} A variety of these fluorescent nucleic acid analogues are depicted in Figure 1 - 9 ((d) - (f)).

![Figure 1 - 9: Various internal fluorescent analogues. (a) 2-AP; (b) dG Pteridine derivatives; (c) dA Pteridine derivatives; (d) – (f) polycyclic hydrocarbon based fluorescent nucleobase derivatives. dR corresponds to the 2'-deoxyribose sugar.]

Most fluorescent nucleobase analogues have a quantum yield that is sensitive to their immediate surroundings. Factors that affect the fluorescence quantum yield of a base analogue include hydrogen bonding, single- or double-strandedness, and neighbouring bases.\textsuperscript{36} Generally, fluorescence of base analogues is highly quenched inside DNA, and normally the effect is most significant if surrounded by neighbouring purines.\textsuperscript{27,31}

When comparing external and internal modifications there seems to be some discrepancies within the literature on which category is “better”. Realistically, depending on the specific application, either external or internal modification could be preferable when using fluorescence to investigate nucleic acid-containing systems. Therefore, it is inappropriate to claim that a certain fluorophore or even group of fluorophores is, overall, best suited for fluorescently modifying DNA. Instead one needs to carefully consider each system under examination and try to optimize the
choice of fluorescent probe molecule thereafter. Externally attached probes can be very bright and photostable and would be best used in investigations where the probe molecule is covalently attached via a long flexible linker and far from the site of interest. In this case, the bulkiness of the external fluorophore could give rise to problems concerning the binding of other molecules to the nucleic acid. Internal modifications such as fluorescent nucleobase analogues can be preferred for investigations where the level of detail needs to be high, since the probe can be designed to be very close to the site of examination. Moreover, the use of nucleobase analogues normally minimizes perturbations to the native structure and behaviour of the nucleic acid. Therefore, it is more likely that fluorescent base analogues generate data that better corresponds to the native behaviour of the system when compared to using external modifications.

Instead of having to generate entirely novel nucleobase analogues, one can employ modifications to the natural nucleobases to dramatically alter their photophysical characteristics. This strategy has been previously employed within the Manderville group and has generated different 8-aryl-dG nucleobases with interesting photophysical properties. These probes have been shown to probe DNA conformation, H-bonding interactions, and metallation. These fluorescent nucleobase analogues are generated by modifying dG and their utilization in numerous nucleic acid systems will be presented and discussed at length throughout this body of work.

1.1.4. DNA Tertiary Structure

1.1.4.1. Duplex DNA

The structure of duplex DNA can assume several distinct forms depending on many factors including, but not limited to solvent composition and base sequence. Biologically speaking, the most prominent global structure DNA adopts is duplex DNA. Duplex DNA has two complementary strands arranged anti-parallel to one another. The duplex DNA structure was solved and reported in 1953 by Watson and Crick. The strands are said to follow complementary base pairing; a scenario
previously discussed involving the ability of a purine nucleobase to bind specifically to its pyrimidine counterpart, such that A pairs with T, and G pairs with C.

B-DNA is the most common structure for duplex DNA; however, there are also other forms, namely A-DNA and Z-DNA. B-DNA is the classical three-dimensional model of the duplex, consisting of two anti-parallel polynucleotide strands winding in a right handed fashion around a common axis (Figure 1 - 10). This positions the hydrophilic backbone on the exterior of the double helix, exposing it to an aqueous environment. The hydrophobic purine and pyrimidine bases are stacked inside the double helix, with the interior of the helix determined to possess a dielectric constant of 3-5. The pairing of the complementary strands results in a major and minor groove on the exterior of the duplex. H-bonding and base stacking interactions account for the stabilization of the duplex through various Van der Waals and dipole-dipole interactions.44

![Figure 1 - 10: Structure of B-DNA viewed from the side (left) and top (right). Hydrophilic phosphate backbones are colored blue while the hydrophobic purine and pyrimidine bases are colored red.](image)

The ideal B-DNA duplex is about 20 Å in diameter, has ten base pairs per turn with a rise per base pair of approximately 3.4 Å. The plane of the base pairs tilts ~2° with respect to the helical axis and has a helical twist of 36°. In B-DNA, the favoured sugar pucker is 2′-endo and the glycosyl
bond is in the *anti* conformation. The other variations of duplex DNA (i.e. A-form and Z-form DNA) typically form under either dehydrating conditions or high salt conditions.  

### 1.1.4.2. Alternative DNA Structures

DNA’s structure is not restricted to duplexes, changes in base sequence and environment can have an enormous influence on DNA structure. DNA’s structural variability is made possible through modifications in the sugar pucker, sugar-phosphate backbone geometry, changes in H-bonding arrangement and the orientation about the glycosyl bond; this allows completely distinct DNA structures to be established. These DNA structures can differ in features such as handedness, base-pairing or the number of strands required for assembly.  

Aside from the previously mentioned A-DNA and Z-DNA other non-B-form secondary structures include G-quadruplex structures (G4 structures) as well as cruciforms and triplexes. These structures were originally characterized *in vitro* using biophysical techniques (for example CD). Accumulating evidence now points towards the existence of these structures under physiologically relevant conditions, and all of them are hypothesized, or even known, to have functional roles *in vivo*.  

“The current wealth of genomic data — which is enabling the evolutionary comparison of motifs that can adopt non-B-form secondary structures *in vitro* — and the use of structure-specific antibodies, structure-binding ligands and clever experimental techniques are driving progress in this field.”

Interest in G-quadruplex structures has increased enormously in recent years, owing to their unique physical properties and the presence of G-rich sequences in biologically functional regions of many genomes. Sequences with G-rich regions able to form G-quadruplexes are over-represented in telomeres, mitotic and meiotic double strand break sites, and transcriptional start sites.
1.1.4.3. Overview of G4 DNA

It is said, and rightfully so, that the story of G-quadruplex DNA started with a Bang. In 1910, Ivar Bang found that guanine behaved differently from all other nucleobases. He found that polyguanylic acid formed a gel at high concentrations, suggesting that a higher-order structure existed. 50 years later the structure was discovered by Gellert and colleagues using X-ray diffraction to demonstrate that guanylic acids can assemble into tetrameric structures. In these tetramers, four guanine molecules form a square planar arrangement in which each guanine is hydrogen bonded to the two adjacent guanines using its W-C, and Hoogsteen H-bonding faces (Figure 1 - 11). The assembly of stacked G-quartets forms the quadruplex structure, and the intervening sequences are extruded as single strand loops. The sequence and size of the loop regions varies, but they typically have less than seven nucleotides, as smaller loops result in more stable G-quadruplex structures. The structure is stabilized by monovalent cations that occupy the central cavities between the stacks, neutralizing the electrostatic repulsion of inwardly pointing guanine O6 oxygens.

Figure 1 - 11: Structure of a G-tetrad which can stack vertically to assemble G quadruplexes.
G-quadruplex structures adopt a variety of topologies, and can be classified into various groups depending on the orientation of the DNA strands (Figure 1 - 9). Structures can be parallel (in which all G’s in the tetrad are anti), anti-parallel (alternating syn and anti G’s) or hybrids thereof. They are also able to form within one strand (intramolecular), or from multiple strands (intermolecular), and various loop structures are possible. Stability of these structures depends on many factors including the length and sequence composition of the total motif, the size of the loops between the guanines, strand stoichiometry and alignment, and the nature of the binding cations.

Figure 1 - 12: Various strand alignments in the formation of G quadruplexes. Quadruplexes can be formed intra- or inter-molecularly and can be comprised of single or multiple (up to four) oligonucleotide strands
1.1.5. Analysis of DNA Structure

1.1.5.1. Thermal Melting

Thermal melting ($T_m$) is defined as temperature of melting, or more appropriately as temperature of midtransition.\textsuperscript{55} A thermal denaturation experiment determines the stability of the secondary structure of DNA or RNA and aids in the determination of the sequence’s ability to form and maintain its folded topology. Experimentally, heating a sample often leads to a change in absorbance properties, which reflects a conformational change of the molecule(s) in solution. With duplex DNA, denaturation results in an increase in the UV absorption at 260 nm and which is called the hyperchromic effect; renaturation accordingly causes a hypochromic effect. In the case of proteins, the denaturation is usually irreversible; however, nucleic acids are able to renature upon cooling.

The formation of nucleic acid duplexes has been studied in great detail, and the parameters that influence duplex stability are relatively well defined. Predictive algorithms to calculate the stability of DNA-DNA, RNA-DNA, or RNA-RNA duplexes exist;\textsuperscript{56,57} however, these rules are empirical, and the actual secondary structure of a nucleic acid may differ significantly from the predicted folding pattern. For instance, one can observe formation of double-stranded structures that do not resemble classic B-DNA by twisting and turning of the nucleic acid to orient the strands in parallel and/or Hoogsteen duplexes.\textsuperscript{58} As previously mentioned, unusual DNA structures such as triplexes and quadruplexes may be implicated in fundamental processes, such as gene expression, and they represent unique targets for both structure-specific and sequence-specific agents.

Other parameters will influence the $T_m$ of the sample. $T_m$ will vary based on the ionic strength and nature of the salt as the cation plays an essential role in stabilization of DNA structures. Cations are concentrated in the vicinity of nucleic acids, and partially neutralize the negative charges.
of their backbone phosphates. A 10 fold increase in ionic strength leads to a $T_m$ increase of 16-18°C for duplexes, which plateaus around 1 M salt concentration.\textsuperscript{59}

1.1.5.2. Circular Dichroism

Circular dichroism (CD) belongs to a group of experimental techniques known as chiroptical methods and has the unique ability to optically differentiate the secondary structures of biological molecules.\textsuperscript{60} The method is based on the interaction between an incident beam of circularly polarized light with either a chiral center in a molecule or chirality imparted by the conformation of a macromolecule. CD requires that a chromophore is present in the molecule for it to be CD active. The physical phenomenon of CD arises from the differential absorption of left- and right-polarized light by a chromophore. If two or more strongly absorbing chromophores are chirally oriented with respect to each other, one observes an excitation spectrum characterized by the presence of two bands with opposite signs, where $\lambda_{\text{max}}$ in absorption corresponds, or nearly corresponds, to zero CD intensity.\textsuperscript{61} The quantity used to describe this phenomenon is called ellipticity, $\theta$, and is expressed in degrees.

The CD of DNA arises from the asymmetric backbone sugars and the helical arrangement of its constituents.\textsuperscript{62} Typically, the CD of electronic transitions of bases within the 200 to 320 nm range is monitored; this reflects the nucleobases’ mutual positions within the duplex.\textsuperscript{46} CD spectroscopy is ideal for tracing conformational transitions between discrete nucleic acid rearrangements; the method is fast, highly sensitive and relatively inexpensive.

The theoretical description of CD spectra of molecules as large as DNA is very complex; therefore, the method is primarily used for empirical studies with DNA. CD offers some advantages over high-resolution methods such as X-ray diffraction of crystals or NMR spectroscopy, which can require crystallization, larger amounts of material, longer accumulation, limited experimental conditions, shorter oligonucleotides and a single discrete structure (as opposed to a mixture of
isomerising conformers). In contrast, CD spectroscopy has a very high degree of sensitivity, permitting work with DNA as low as 20 μg/mL. The studied molecules can be either oligonucleotides or polynucleotides, which is especially important with DNA, and CD can be used to distinguish conformational isomerizations between distinct conformers from gradual changes within their arrangements. The prevalence of CD data in the literature speaks to its applicability, as practically any key discovery regarding secondary structures of DNA include CD characterization.

Of the numerous forms of DNA discussed up to this point, duplex DNA and quadruplex DNA will be interpreted via CD spectroscopy, and will be the focus of this section. Generic spectra of various duplex DNAs are presented in Figure 1 - 13. Distinguishing between the possible duplexes is possible, B-DNA provides a CD spectrum with small amplitude bands including a positive band around 280 nm and a negative one around 245 nm, with a crossover at approximately 260 nm. A-DNA exhibits large positive (270 nm) and small negative (230 nm) bands, with a crossover at 240 nm, while Z-form DNA exhibits small positive (285 nm) and large negative (245 nm) with a crossover at 280 nm. Depending on the base sequence, DNA molecules adopt more or less different structures, and thus, CD spectra differ as well.

Figure 1 - 13: Characteristic CD spectra for B-, Z-, and A-form DNA. Reproduced from: http://homepage.univie.ac.at/johannes.winkler/cd.html
CD spectroscopy can also discriminate between the particular types of quadruplex folding (Figure 1 - 14).\(^{46}\) Parallel quadruplexes are characterized by a dominant positive band at 260 nm and a negative band at 240 nm, whereas the spectra of anti-parallel forms display the inverse of the parallel form (negative at 260 nm and positive at 240 nm) plus an additional positive band at 295 nm. In addition, most quadruplexes display another positive peak around 215 nm.

![CD spectra of guanine quadruplexes. Left side: the parallel stranded quadruplex \([d(G_4)]_4\) stabilized by 16 mM K\(^+\); right side: Na\(^+\)-induced anti-parallel bimolecular quadruplex of \([d(G_4T_4G_4)]_2\). The triangles in the sketches indicate guanines and point in the 5′–3′ direction. Reproduced from reference with permission.\(^{46}\)](image)

The differences between the CD spectra of the various quadruplex forms have been attributed to the stacking interactions of neighboring G-quartets with the same or opposite polarity.\(^{61}\) This polarity can be represented by designating one side of the quartet the “head” while the other is
the “tail.” By convention, the “head” side refers to side of the quartet that has the donor to acceptor H-bonding running clockwise, whereas on the “tail” side the donor to acceptor H-bonding interaction runs counter-clockwise, as can be seen in Figure 1 - 15. The stacking arrangements of the adjacent quartets in either a head to head, head to tail or tail to tail orientation generate the different CD features irrespective of the relative orientation of the strands (parallel/anti-parallel).61

Figure 1 - 15: The G-Quartet (a) shown on its “head” (H) side (from donor to acceptor H-bonding runs clockwise); (b) The “tail” side (donor from acceptor H-bonding runs counter clockwise); (c) Heteropolar stacking arrangement of two G-tetrads (head to tail stacked with one another); (d) Homopolar stacking arrangement of two G-tetrads (head to head stacked with one another).
When the glycosyl bonds of the guanines alternate in \textit{syn} and \textit{anti} conformations along each strand, as is the case for anti-parallel quadruplexes, the G-quartet polarity also alternates. This results in heteropolar stacking (Figure 1 - 15 (c)) and this spectrum would resemble that of Figure 1 - 14 (a). When the glycosyl bonds of the guanines are all \textit{anti}, G-quartet polarity is the same and stacked in a homopolar arrangement; here the CD spectra would resemble Figure 1 - 14 (b).

\textbf{1.1.6. DNA Damage}

\textbf{1.1.6.1. Mutagenesis}

DNA damage plays a major role in mutagenesis, carcinogenesis and aging. The chemical events leading to DNA damage include: hydrolysis, exposure to reactive oxygen species (ROS) and other reactive metabolites leading to adducts. These reactions are triggered by exposure to exogenous chemicals, or can result from metabolic, endogenous processes.\textsuperscript{63} Examples of exogenous chemicals include: alcohol, tobacco, asbestos and air pollution.\textsuperscript{64} However, one cannot simply attribute the incidence of cancers present in our population to the concentrations and mutagenic potentials of known carcinogens. Mutations due to DNA damage, caused by unidentified exogenous agents, and to an increase in endogenous damage which is mediated by exogenous factors (such as depletion of radical scavengers) must play a role in most cases of cancer.\textsuperscript{65} Endogenous DNA damage occurs at a high frequency and the types of damage produced by normal cellular processes are identical or very similar to those caused by some environmental agents.

Mutations refer to a change of the nucleotide sequence of the genome of an organism. Examples of DNA mutations include point mutations or single nucleotide polymorphisms (SNPs), insertions, or deletions. SNPs involve the replacement of a nucleotide with another. If these mutations are termed transversions, the situation involves a purine being substituted by a pyrimidine, or vice versa. Alternatively, transitions describe a situation where a purine is substituted with another purine, or pyrimidine substituted with another pyrimidine.\textsuperscript{44} Insertion mutations involve the

25
addition of one or more nucleotides or base pairs into a DNA sequence; as opposed to a deletion mutation which describes the loss of one or more nucleotides or base pairs. These types of mutations typically give rise to frame shifts, which results when a given number of nucleotides are not divisible by three. During translation, genes are interpreted as a series of codons that each have three nucleotides. Inserting or deleting a number of bases that is not divisible by three will shift the frame of translation, resulting in a completely different protein sequence. One can imagine that the earlier in the sequence the deletion or insertion occurs, the more deleterious it is on the translation of the sequence.

1.1.6.2. Oxidative DNA damage

In living cells reactive oxygen species (ROS) are formed continuously as a consequence of metabolic and other biochemical reactions, as well as external factors. Oxidative metabolism is essential to life, as the oxidation of organic compounds allows for the release of free energy required for many life functions. DNA damage produced by endogenous sources is considered to be the most frequently occurring type of damage. Exposure to such ROS can lead to several types of DNA damage including oxidized bases and/or single- and double-strand breaks. The most common ROS and their sources are listed in Table 1 - 3, these include superoxide (O$_2^-$), hydrogen peroxide (H$_2$O$_2$), hydroxyl radicals (OH$^-$) and singlet oxygen (¹O$_2$).
Table 1 - 3: Sources of reactive oxygen species

<table>
<thead>
<tr>
<th>Cellular Oxidants</th>
<th>Source</th>
<th>Oxidative Species</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endogenous</strong></td>
<td>Mitochondria</td>
<td>( \text{O}_2^-, \text{H}_2\text{O}_2^-, \text{OH} )</td>
</tr>
<tr>
<td></td>
<td>Cytochrome P450</td>
<td>( \text{O}_2^-, \text{H}_2\text{O}_2 )</td>
</tr>
<tr>
<td></td>
<td>Macrophage cells</td>
<td>( \text{O}_2^-, \text{H}_2\text{O}_2 )</td>
</tr>
<tr>
<td></td>
<td>Peroxisomes</td>
<td>( \text{H}_2\text{O}_2 )</td>
</tr>
<tr>
<td><strong>Exogenous</strong></td>
<td>Redox cycling compounds</td>
<td>( \text{O}_2^- )</td>
</tr>
<tr>
<td></td>
<td>Metals (Fenton reaction)</td>
<td>( \cdot\text{OH} )</td>
</tr>
<tr>
<td></td>
<td>Radiation</td>
<td>( \cdot\text{OH} )</td>
</tr>
</tbody>
</table>

During mitochondrial oxidative metabolism, the majority of the oxygen consumed is reduced to water; however, an estimated 4% to 5% of molecular oxygen is converted to ROS, primarily superoxide anion, formed by an initial one-electron reduction of molecular oxygen. Superoxide can be dismutated by superoxide dismutase to yield \( \text{H}_2\text{O}_2 \). In the presence of partially-reduced metal ions, particularly \( \text{Fe}(II) \), hydrogen peroxide is subsequently converted through Fenton and Haber-Weiss reactions to a hydroxyl radical. The hydroxyl radical is highly reactive and can interact with nucleic acids, lipids, and proteins.

Hydroxyl radical is known to react directly with \( \text{dG} \) to form 7,8-dihydro-8-oxo-2′-deoxyguanosine (8-oxo-dG), shown in Figure 1 - 17. This is the most common oxidative lesion (owing to \( \text{dG} \) having the lowest oxidation potential of the nucleobases) observed in duplex DNA and subsequently serves as a biomarker for cellular oxidative damage. Numerous studies have focused on misreading and mutations induced by 8-oxo-dG. Its ability to form both a W-C base pair with \( \text{dC} \) and a Hoogsteen-like base pair with \( \text{dA} \) leads to replicative errors resulting in \( \text{G} \rightarrow \text{T} \) transversion mutations. Moreover, once 8-oxo-dG has been formed, it is more easily oxidized than any of the normal nucleobases, with an oxidation potential of 0.74 V/NHE. This leads to the
production of hydantoin lesions of 8-oxo-dG, which have been shown to be more detrimental to the genome than 8-oxo-dG itself, causing several other transversion mutations.\textsuperscript{72-74}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{lesions.png}
\caption{The structure of 8-oxo-G, guanidinohydantoin heterocycle (Gh) and spiroiminodihydantoin (Sp) lesions.}
\end{figure}

\subsection{1.1.6.3. Hydrolysis of Nucleotides}

Hydrolysis of nucleotides is a process in which the C1'-N9 glycosidic bond between the nucleobase and deoxyribose sugar is broken, and can occur spontaneously. The result is a loss of the nucleobase from the oligonucleotide, generating what is called an abasic site. Depurination occurs at a much higher rate than depyrimidination.\textsuperscript{75,76} Under typical cellular conditions, approximately one in $10^5$ purines is lost from DNA every day.\textsuperscript{7} Another common form of damage is deamination, caused by hydrolysis of the nucleobases exocyclic amino groups; this situation occurs in about one of every $10^7$ nucleotides per day under normal cellular conditions.\textsuperscript{7} Hydrolysis of various modified dG nucleosides will be discussed in detail in Chapter 2.

\subsection{1.1.6.4. Adduct Formation}

Evidence that chemicals cause cancer has long been known. Physicians in the 18\textsuperscript{th} century associated nasal and oral tumors with the use of snuff tobacco, and subsequently, Pott associated scrotal tumors in chimney sweeps with the soot and tar the boys were exposed to.\textsuperscript{77}
Extensive efforts beginning in the 1940’s have shown that many chemicals must be converted to reactive forms in the body in order to cause cancer. DNA is nucleophilic in nature, therefore, each nucleobase is vulnerable to covalent modification by electrophiles at various positions to form adducts. Reactive metabolites in the body attach to DNA and proteins, and the structures of many of these DNA carcinogen adducts have now been characterized.

The pattern of DNA adduct formation is not only highly dependent on the reactive intermediate involved, but also on the nucleobase sites available for modification, as electrophiles are site-specific in their reactivity. The target sites for covalent modification are nucleophilic atoms with lone pairs of electrons or a negative charge. ‘Hard’ electrophiles, such as carbenium and nitrenium ions, prefer to react with similarly ‘hard’ nucleophilic centers, such as the lone pairs of electrons on the amino nitrogen and keto oxygen atoms of the nucleobases. The pattern of adduct formation also differs greatly within duplex DNA, due in part to the steric and electronic environment of dsDNA.

Depicted below in Figure 1 - 17 are the sites prone to reactions with electrophiles for dG. As mentioned earlier, due to dG’s electron-rich character and oxidation potential, it is the most commonly targeted nucleobase for attachment of electrophilic species.

![Chemical structures and reactions](image)

**Figure 1 - 17**: Site-specific modifications of dG by chemical carcinogens.
Chemical carcinogens that transfer alkyl moieties typically react with the endocyclic N7 and exocyclic O6 atoms. This type of DNA damage has been employed by common warfare agents such as mustard gas. The $N^2$ and $N^4$ sites are both able to be modified by Michael acceptors and quinone methides, and the $N^2$ atom is also targeted by polycyclic aromatic hydrocarbons (PAHs). PAHs are ubiquitous environmental pollutants produced in the combustion of organic matter,\textsuperscript{80,81} for example in furnaces, gasoline/diesel engines and tobacco smoke.\textsuperscript{80} The C8-site of dG is prone to attack by aryl radicals.\textsuperscript{82,83} Typically, such aryl radicals are derived from aromatic and heterocyclic amines, nitroaromatics, PAHs, and small aryl and phenol moieties.\textsuperscript{83}

1.1.6.5. C8-Aryl-Modified dG Nucleosides

C8-aryl-dG adducts are classified as either O-linked, N-linked or C-linked referring to the atom that is directly attached to the C8 site. N-Linked C8-aryl-dG adducts have been extensively studied; they are derived from aromatic amines and their nitro derivatives.\textsuperscript{84} Of the plethora of N-linked adducts studied, the most attention has been paid to C8- and N-2-amino-3-methylimidazo[4,5-f]quinoline (IQ-dG),\textsuperscript{85} N-2-aminofluorene (AF-dG),\textsuperscript{12,86} N-acetyl-2-aminofluorene (AAF-dG)\textsuperscript{87} and N-2-fluoroaminofluorene (FAF-dG)\textsuperscript{88,89} as well as the N-(deoxyguanosine-8-yl)-1-aminopyrene (AP-dG or dG\textsuperscript{AP}) (Figure 1 - 18).\textsuperscript{90} Despite the structural similarities within the family of AF-dG, AAF-dG and FAF-dG, these adducts have been found to exhibit unique mutation and repair activities.\textsuperscript{84}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Structures of commonly studied N-linked dG DNA adducts.}
\end{figure}
The distinct mutagenic properties of the acetylated and non-acetylated aromatic amine lesions are caused by their different conformational preferences. These conformational preferences result in sequence-dependent equilibrium between the external (B-type), stacked (S-type) and wedge (W-type) conformations within fully paired DNA duplexes (Figure 1 - 19). These have been extensively characterized by fluorescence, \(^{89}\) \(^{1}\)H NMR, \(^{91}\) crystallographic analysis, \(^{92}\) \(^{19}\)F NMR, \(^{14}\) as well as CD.\(^{12}\) More specific information pertaining to this family of adducts will be presented in Chapter 3.

Modification at the C8 site of dG significantly impacts the \(antil/syn\) conformational equilibrium, which impacts the nucleosides structure within duplex DNA. In general, within the duplex, there exist three conformations (Figure 1 - 19). The least perturbing conformation is the "B-type" (B) in which the modified guanine residue retains its \(anti\) conformation, as well as its W-C H-bonding interaction with the opposing strand; this conformation projects the adducted moiety into the major groove and exposes it to an aqueous environment. The base-displaced "stacked" (S) conformer involves the C8-dG adduct adopting a \(syn\) orientation, disrupting the W-C H-bonding interaction, since it points toward the major groove. In this scenario, the adducted moiety is placed within the interior of the helix, forcing the opposing base to be flipped out of the helix, and the modification stacks with neighbouring bases. The "wedge" (W) conformation typically arises when the C8-dG adduct is mispaired opposite another purine base.\(^{86}\) This scenario involves the modified base adopting a \(syn\) conformation; however, the guanine portion remains stacked within the helix placing the adducted moiety in the minor groove. Studies on a variety of lesions have shown that adducts are capable of forming more than one conformer, which results in a mixture of conformations within the duplex. The three conformers depicted in Figure 1 - 20 have the C8 modification termed “Aryl” for simplicity.
The three major conformers of C8-modified deoxyguanosine adducts within duplex DNA. Aryl is used to denote the adducted moiety.

Figure 1 - 19: The three major conformers of C8-modified deoxyguanosine adducts within duplex DNA. Aryl is used to denote the adducted moiety.

C-Linked adducts have also been studied and predicted to form the same structures as those mentioned above. Molecular dynamics simulations carried out by the Wetmore laboratory concluded that syn and anti structures of both ortho and para C-linked phenolic adducts were similar in energy; thus, more than one conformation within the DNA strand was expected. In addition, CD data from C8-aryl adducts studied showed regular B-form DNA for ortho and para phenol adducts; however, the 8-benzothienyl-dG (Bth-dG) adduct produced an induced CD band, which was interpreted as the syn, S structure. Further investigation into oligonucleotide conformations of phenolic-, furanyl-, and benzothiophenyl-dG modified duplexes indicated that conformational heterogeneity depended on the opposing base. C8-aryl adducts incorporated opposite the correctly
paired C were found to prefer anti-B-type structures, while base pairing to a mismatched G rendered primarily W structures.37,41

1.1.6.6. Toxicological Relevance of C8 Adducts

Incorporation of C8 adducts into oligonucleotides has determined that their effect on DNA structure is modulated by the adducts’ conformational preferences and the flanking sequence context. This polymorphism accounts for a range of toxicological outcomes. N-Linked aromatic amines may adopt either pro-mutagenic or non-mutagenic conformations within the DNA environment. The pro-mutagenic conformations of C8-aryl-amine adducts are believed to be the syn conformations (S and W conformers), since it has been observed that relative mutagenicity parallels the conformational preference for these structures.91

Polymerase studies have identified sequences that are more susceptible to producing frameshift mutations.96 The recognition sequence of the restriction endonuclease (5'-G1G2CG3CC-3') is one such location. This sequence is called the "NarI" recognition sequence named after the bacteria which carries this gene, Nocardia aregentinesis.97 The site contains a GC repeat that is prone to spontaneous and induced frameshift mutations. Modification at the C8 position of the G3 guanine by bulky arylamines is known to induce -2 frameshift mutations.98,99 This phenomenon has been explained using the slippage model first developed by Streisinger (Scheme 1 - 1).100
The slippage model presented in Scheme 1 - 1 presents the findings of Rizzo, who subjected the *NarI* sequence containing IQ-dG to polymerase extension studies with the prokaryotic polymerase, pol II. Replication was found to proceed normally until the modified base (denoted X) was encountered. A dC base was then incorporated opposite the lesion; however, because this base could then normally base-pair with a neighbouring G two bases removed from the lesion in the 5' direction, a slippage occurred, resulting in a bulge. This misalignment of base pairing can happen in regions of base repeats and is stabilized by the bulky IQ-dG adduct. It was found that the S-form duplex conformation of the IQ-dG accounted for stabilization of the bulged out sequence. This results in the appearance of a -2 frameshift mutation, as there was loss of two base pairs from the replicated strand. The propensity of *N*-linked arylamine adducts to form -2 frameshift mutations has been shown to correlate with their conformational preference for the S-form conformation.
In order to test the ability of an adduct to induce frameshift mutations, hybridization to the resulting slipped sequence has been used to determine whether a particular adduct can stabilize bulge formation.  

1.2. Purpose of Research

Research conducted by the Manderville group is motivated by interest of C-linked C8-aryl-dG adducts, derived from phenolic toxins reaction with oligonucleotides. This interest has given rise to the synthesis and study of various C8-aryl substituted dG nucleoside adducts in order to better understand their biological implications.

Through previous work within the Manderville group, the oxidative properties of different C-linked C8-aryl-dG adducts by electrochemical oxidation have been investigated. C-Linked C8-aryl-dG adducts have been found to be more readily oxidized than dG, implicating them in DNA damage. While adduct formation is viewed as ultimately detrimental and a form of DNA damage, C8-aryl-dG adducts possess properties that can be exploited to determine DNA structure and conformation. While the native nucleosides only very weakly fluorescent, modification at the C8-position of dG with an aryl moiety gives a strongly fluorescent molecule. The fluorescent properties of a variety of C8-heteroaryl-dG nucleoside adducts have been studied, and their potential for the utilization to probe various DNA environments is beginning to be realized. Unfortunately, while providing emissive analogues, the C8-aryl-dG adducts are also more prone to undergo depurination, which results in abasic sites. This has led to some difficulties incorporating such adducts into oligonucleotide strands, as typical solid-phase synthesis strategies include multiple exposures to acid, which result in degradation of the oligonucleotide. An alternative method that optimizes the incorporation of these acid sensitive adducts into oligonucleotides was developed and is presented in Chapter 2.

A series of adducts have been incorporated and tested with various oligonucleotide sequences. The results presented will demonstrate that the fluorescence sensitivity exhibited by such
adducts towards their local environment allows for insight into the conformational heterogeneity of the modified nucleoside within dsDNA (Chapter 3). These studies provide structural insight that relates (and contrasts) to persistent N- and C-linked C8-dG adducts formed in vivo from known carcinogens, and which will help to provide further understanding for the mechanisms of adduct-induced mutagenesis. Genuine 22-mer NarI oligonucleotides have been prepared to be biochemically studied using different DNA polymerases to determine how replication of such adducts proceeds.

Chapter 4 will expand on the applications of fluorescent 8-aryl-dG adducts, as the focus will shift to the characterization of various 8-aryl-dG probes within quadruplex DNA. It became evident that the probes were able to distinguish various structural transformations within the oligonucleotide. This led to the realization of their potential use within DNA-based diagnostics for probing duplex–quadruplex exchange; this is a common strategy for target detection using fluorescent probes that turn-on during the exchange process. Here, the utility of internal fluorescent 8-aryl-2’-dG probes for detecting G-quadruplex folding within the thrombin-binding aptamer (TBA) will be demonstrated. Studies will also demonstrate how the choice of 8-aryl substituent can be used to tune probe electronics for turn-on fluorescence in the duplex or G-quadruplex structure. Overall, this study established 8-aryl-dG probes as useful biophysical tools for DNA-based diagnostics. This chapter will also show how 8-aryl-dG adducts can be utilized in aptamer development, which is a fascinating and rapidly growing field.

The diagnostic utility of 8-aryl-dG adducts is then highlighted within the human telomeric sequence (HTelo), which is discussed in chapter 5. Here, the goal was to determine the 8-aryl-dG adduct impact on the complex mixture of topologies adopted by the HTelo sequence. Currently, extensive efforts are aimed at inducing the formation and stabilization of certain quadruplex topologies within this sequence, as it has been found that certain topologies result in decreased telomerase activity. Telomerase is over expressed in over 85% of all cancers. The ability of 8-aryl-
dG probes to distinguish the associated conformational heterogeneity within the HTelo sequence will be presented therein.
1.3. References


(2) Avery, O. T.; Macleod, C. M.; McCarty, M. J. Exp. Med. 1944, 79, 137.


(53) Hardin, C. C.; Perry, A. G.; White, K. Biopolymers 2000, 56, 147.


(64) Harris, P. J.; Ferguson, L. R. *Mutat. Res.* **1999**, *443*, 95.


(87) Zhou, Y.; Romano, L. J. Biochemistry 1993, 32, 14043.


(97) Fuchs, R. P.; Fujii, S. DNA Repair 2007, 6, 1032.


Chapter 2: Solid Phase Synthesis of Oligonucleotides Containing Acid-Sensitive 8-Aryl-Guanine Adducts
Note on Reproduction of Work

Portions of the research presented in this chapter have been previously published as a first-authored journal article, with the corresponding citation provided below, and are reproduced herein with permission from the respective journal.

2.1. BIG PICTURE

In order to study the biological fate of a DNA lesion, one typically synthesizes modified oligonucleotides which act as authentic standards, and subjects them to various polymerases. Aryl radical species derived from the metabolism of polyaromatic hydrocarbons (PAHs),\textsuperscript{1} arylhydrazines,\textsuperscript{2,3} estrogens\textsuperscript{4} and phenolic toxins\textsuperscript{5,6} can attach to the 8-position of 2'-deoxyguanosine (dG) resulting in carbon-linked 8-aryl-dG adducts. If left unrepaired, these lesions may be carcinogenic. The 8-aryl-dG adducts are also highly emissive and are useful fluorescent nucleobase probes for detecting G-quadruplex folding\textsuperscript{5} and for monitoring adduct conformation within duplex DNA.\textsuperscript{7}

To study the structural, photophysical and biological properties of 8-aryl-dG adducts in oligonucleotide substrates, we aim to incorporate them into the G$_3$-site of the 12-mer sequence 5'-CTCGGCXCCATC (X = 8-aryl-dG), which contains the recognition sequence of the NarI Type II restriction endonuclease. This sequence represents a “hotspot” for mutagenicity mediated by nitrogen-linked 8-aryl-dG adducts produced by arylamine carcinogens.\textsuperscript{8} The 12-mer sequence permits analysis of adduct impact on duplex DNA structure; for N-linked arylamine adducts, such duplex structures have been used to predict mutagenic outcome,\textsuperscript{8} and propensity for DNA repair.\textsuperscript{9,10} A second goal is to incorporate 8-aryl-dG adducts into longer DNA substrates (NarI(22)) that can be used for primer-extension assays using DNA polymerases to assess the biological impact of these lesions.

In an effort to achieve these goals and generate the necessary oligonucleotide substrates containing 8-aryl-dG lesions, the Manderville group has previously utilized a post-synthetic strategy using Suzuki cross-coupling,\textsuperscript{11} which avoids synthesis of the modified phosphoramidite and exposure of the 8-aryl-dG lesion to acidic deblock (3% dichloroacetic acid (DCA) for removal of the 5'-O-dimethoxytrityl (DMT) protecting group). This approach can be used to incorporate a single adduct into relatively short substrates (3-15 mers).\textsuperscript{6,12,13} However, yields can
be low and the strategy becomes problematic when either incorporating multiple adducts into strands, or generating longer modified DNAs required to assess the biological impact of the lesion using DNA polymerases and DNA repair enzymes.

The limitations of the Suzuki cross-coupling approach prompted efforts to seek a solid-phase synthesis strategy for incorporating 8-aryl-dG adducts into oligonucleotide substrates. Inspired by recent efforts of the Yan laboratory, we deemed the 2,7-dimethylpixyl (DMPx) protecting group as the most attractive alternative to the widely used 5′-O-DMT. The DMPx group is more acid-labile than DMT and can be efficiently removed using 0.5% DCA in dry dichloromethane (DCM). This suggested that the DMPx group may offer a general protocol for efficient solid-phase synthesis of oligonucleotides containing acid-sensitive 8-aryl-dG adducts. This chapter outlines our efforts to optimize solid-phase synthesis of NarI substrates containing 8-aryl-dG adducts and demonstrate the utility of the 5′-O-DMPx protecting group for efficient synthesis of oligonucleotide substrates containing acid-sensitive 8-aryl-dG adducts.

2.2. General Introduction

Unmodified DNA bases are extremely stable to hydrolysis under physiological conditions; so stable indeed that half-lives of hydrolysis have been reported to be 730 years for the purines, and 14,700 years for the pyrimidines. However, nucleobase hydrolysis rates increase significantly in the presence of acid, heat, methylation or by the action of enzymes. Hydrolysis of the glycosidic bond is generally termed deglycosylation; since purines are much more sensitive than pyrimidines to deglycosylation, the process is sometimes more specifically referred to as depurination. This results in the generation of the free purine base, and an abasic, or apurinic (AP) site (shown in Scheme 2 - 1 for the hydrolysis of dG within DNA). The abasic site may exist in three different forms at equilibrium, namely, the open chain aldehyde, the hemiacetals and hydrate. Although the aldehyde form constitutes only 1% of the mixture at
equilibrium, which is dominated by the hemiacetals, the aldehyde form confers to the abasic site’s high sensitivity to alkaline conditions. Typically, abasic sites undergo β-elimination under alkaline conditions to form the α-β unsaturated aldehyde.\textsuperscript{17} Under more strongly alkaline conditions, a second elimination (δ elimination) occurs.\textsuperscript{15} This also occurs in the presence of polyamines.\textsuperscript{18}

\begin{center}
Scheme 2 - 1: Deglycosylation (depurination) of a dG nucleobase within DNA and subsequent base-catalyzed strand scission.
\end{center}

Apurinic sites are mutagenic.\textsuperscript{19} Mutations can arise through different mechanisms. One scenario, originating from depurination of a dG residue to afford an abasic site, involves an error prone DNA repair mechanism attempting to repair the lesion which results in G → T transversions.\textsuperscript{20} This repair mechanism, albeit error-prone, gives the cell a way to avoid an AP site being encountered during DNA replication which results in much more severely mutagenic +1 or -1 frameshift mutations.\textsuperscript{21} If AP sites persist unrepaired, they can also undergo spontaneous strand cleavage through β-elimination; this process has a half-life of \textasciitilde 200 hours in pH 7.4 at 37 °C.\textsuperscript{13}
The acid catalyzed hydrolysis of dG is shown in Scheme 2. Under acidic conditions the half-life for dG is accelerated from 730 years at pH 7.4, 37 °C to 17.7 minutes at pH 1, 37 °C. The reaction proceeds via a stepwise mechanism involving a pre-equilibrium protonation step at N7. This protonation is defined by the acid dissociation constant ($K_{a1}$), and precedes the unimolecular rate-limiting cleavage of the glycosidic bond for the mono-protonated species, defined by the rate constant ($k_1$). The $pK_a$ value for N7-protonated dG was determined by Sigel and co-workers to be 2.34. Protonation at N7 to form the mono-cation ($K_{a1}$) can be followed by protonation at N3 to form the di-cation under very acidic conditions ($pK_{a2} = -2.5$) and both substrates are subject to glycosidic bond cleavage since the positively charged purine ring is a good leaving group. This S_N1 type mechanism is formally classified under IUPAC nomenclature as a DN*AN reaction. This notation is interpreted as the leaving group dissociation “DN” (the guanine nucleobase) is being followed by nucleophilic addition “AN” (water) with a short lived intermediate “*” (oxocarbenium ion).

A well established property of C8-aryl-adducts is their tendency to undergo depurination in studies carried out at physiological pH. When comparing 8-aryl-dG adducts to dG, it has been found that the presence of a C8-substituent does not significantly alter $K_{a1}$. Substitution at the
C8-site of dG with electron-withdrawing NO$_2$\textsuperscript{27,28} or SO$_2$CH$_3$\textsuperscript{29} moieties greatly accelerates hydrolysis, while electron-donating NH$_2$\textsuperscript{30} and OCH$_3$\textsuperscript{31} C8-substituents have been shown to decrease the rate of hydrolysis. Interestingly, bulky arylamino\textsuperscript{26} and dimethylamino\textsuperscript{30} C8-substituents have been found to accelerate the rate of hydrolysis despite their electron-donating character. This effect has been attributed to the release of steric strain that occurs upon removal of the deoxyribose moiety.\textsuperscript{26,32}

Studies have established the hydrolytic stability of a class of substituted benzene C8-aryl adducts in 0.1 M HCl at 37 °C.\textsuperscript{33,34} The modified bases studied therein hydrolyzed 5- to 45-fold faster than dG. Adducts with electron-withdrawing substituents were the most labile, while adducts possessing electron donating substituents were the most stable. These findings demonstrated that C8aryl-dG nucleoside adducts were more prone than dG to acid catalyzed hydrolysis. Interestingly, the C8-aryl-dG adducts bearing \textit{para} substituents had $k_1$ values that were 90-200 fold larger than that for dG, while the effects of \textit{ortho} substituted adducts were only 9-60 fold larger. Differences in rates were again attributed to relief of steric strain, upon removal of the deoxyribose sugar moiety.

As seen in Figure 2 - 1, \textit{ortho} and \textit{para} substituted nucleoside derivatives differ in their ability to attain planar nucleobases upon deglycosylation. \textit{Para}-substituted C8-aryl-dG adducts undergo a decrease in twist angle of approximately 40° upon removal of the sugar group from the twisted N$^7$-protonated species to give the planar nucleobase. \textit{Ortho} substituent’s that were able to form H-bonding interactions with the N$^7$-protonated nucleoside, were not as twisted initially, which resulted in smaller decreases in twist angle upon sugar removal to form the planar nucleobase. This resulted in a lower degree of relief in steric strain for these adducts upon sugar removal, which was observed as a slower rate of hydrolysis. \textit{Ortho} substituted derivatives that could not H-bond with the protonated N$^7$ nucleoside were twisted to a higher extent initially, but deglycosylation did not afford a planar, conjugated nucleobase due to steric interactions. Rankin attributed the lack of steric strain relief resulting in a smaller driving force for deglycosylation.
Additionally, one would expect that the ortho substituted adducts’ inability to obtain a planar orientation would result in lower degrees of π-orbital overlap; this lowers the amount of stability instilled through conjugation upon sugar loss and would also result in slower rates of hydrolysis. Despite the enhanced reactivity of 8-aryl-dG adducts in acid compared to dG, Rankin found that they were relatively stable at physiological pH with $t_{1/2} \sim 25$ days.\textsuperscript{34}

![Diagram of substituted benzene C8-aryl adducts](image)

Figure 2 - 1: Deglycosylation and associated half-lives of hydrolysis of substituted benzene C8-aryl adducts driven by the relief of steric strain upon sugar loss.

C8-Aryl-dG nucleoside adducts were also incorporated into trimers for the determination of rates for hydrolysis within DNA.\textsuperscript{34} C8-aryl-dG adducts were incorporated into varying trimer substrates (5’-NXN; N refers to a normal nucleoside while X is the C8-aryl-dG adduct), by application of a post-synthetic method involving palladium-catalyzed Suzuki Miyaura cross-coupling reactions with a brominated trimer, 5’-N(8-Br-G)N, and 10 equivalents of aryl boronic
acid, to give the modified trimers. The expectation was that, upon incorporation within DNA, C8-aryl-dG adducts would exhibit slower rates of hydrolysis. The hydrolysis of the modified trimers was measured by fluorescence spectroscopy, by monitoring the appearance of the deglycosylated product at its excitation and emission maxima. First-order rate constants and half-lives were determined for the modified trimers in 0.1 M HCl at 37.2°C showing that the trimers underwent hydrolysis approximately 25 times slower than the modified nucleosides. The trimers’ rates of hydrolysis also paralleled the trends for the nucleobase adducts, as electron-withdrawing groups exhibited the largest rates of hydrolysis and electron donating groups exhibited the smallest.

Various methods are available for incorporation of modified nucleobases into DNA as will be discussed later. Rankin’s decision to use a Suzuki-Miyaura post-synthetic approach stemmed largely from the fact that C8-aryl-dG adducts are highly sensitive to acid and oxidation. Since the C-linked C8-aryl-dG adducts studied by the Manderville group are commonly synthesized by Suzuki-Miyaura coupling reactions, it seemed as though this methodology would be applicable for the synthesis of C8-aryl-purine modified oligonucleotides. The Pd-catalyzed cross-coupling reaction is noted for its wide tolerance toward a variety of functionalities and applicability to extremely labile systems such as nucleoside monophosphates and triphosphates.

Mechanistically a Suzuki-Miyaura cross coupling reaction involves three major steps in a catalytic cycle (depicted in Scheme 2 - 3). A ligand (L), usually triphenylphosphine or a similar derivative, activates the palladium catalyst. In the first step of the cycle, the palladium catalyst couples to the halide (X) via oxidative addition to form the organopalladium complex, R-Pd(II)L₂-X. Transmetallation with organoboron compound occurs next, forming the organopalladium complex R-Pd(II)L₂-R’. Typically, a base is required to activate the organoboron compound although the mechanism of the transmetallation step remains in question.
Finally, reductive elimination is carried out to give the desired product and regenerate the palladium catalyst for reuse in the catalytic cycle.\textsuperscript{35}

Scheme 2 - 3: Catalytic cycle of the Suzuki-Miyuara cross-coupling reaction.

Suzuki-Miyuara cross-couplings for the synthesis of the C\textsuperscript{8}-aryl-dG adducts have been reported extensively in the literature, and are based on a procedure designed by Shaughnessy and coworkers.\textsuperscript{36,37} These reactions use an aryl-boronic acid (Ar-B(OH)\textsubscript{2}) as the organoboron derivative and 8-bromo-2'-deoxyguanosine (8-Br-dG) as the organic halide. Bromination of the C\textsuperscript{8}-site of dG by reaction with N-bromosuccinimide (NBS), as shown in Figure 2 - 2, provides an efficient generation of starting material. The reaction is carried out in the presence of a water-soluble catalyst, palladium (II) acetate (Pd(OAc)\textsubscript{2}) and the phosphine ligand, tris(3-sulfophenyl)phosphine trisodium salt hydrate (TPPTS).\textsuperscript{37} Rankin extended this methodology to
DNA using a brominated dG trimer, which was purchased for generation of the 8-aryl-dG modified oligonucleotides.

![Chemical Reaction Diagram]

Figure 2 - 2: General synthesis of 8-aryl-dG adducts using 8-Br-dG and the Suzuki-Miyaura coupling methodology.

2.2.1. Incorporation of Modified Nucleosides into DNA

A number of methods for incorporating modified nucleobases into oligonucleotides have been developed and utilized; the four most common approaches are depicted in Figure 2 - 3.\textsuperscript{38} The generation of oligonucleotides containing modified nucleobases allows researchers to determine the fate of such chemical lesions,\textsuperscript{39} and site-specifically modified oligonucleotides are extremely useful probes for examining the mechanism of mutagenesis and carcinogenesis.\textsuperscript{40}
The modified deoxyribonucleosides can be prepared and incorporated using either standard or modified phosphoramidite oligonucleotide synthesis (Figure 2–3 (A)). The modified base must be compatible with the reagents used for preparation of the deoxynucleoside phosphoramidite and subsequent deprotection. The phosphoramidite approach has become the industry standard and is used with commercial solid-phase methods. Once prepared, the phosphoramidites can be incorporated into oligonucleotides through the use of a DNA synthesizer, as discussed in the next section.

The second method for synthesizing oligonucleotides containing DNA adducts at defined positions involves treating an oligonucleotide with a reagent to modify the desired residue (Figure 2–3 (B)). This strategy is useful when the chemistry is incompatible with phosphoramidite and deprotection chemistry, or when the electrophile is specific for reaction with a certain nucleobase. One of the inherent problems with this technique is that multiple modifications may occur. In some cases positional isomers can be separated by chromatography. If this approach is used, the
researcher must bear the burden of rigorously documenting the identity and purity of the products, as the potential for artefacts is very high.

Another approach to synthesize oligonucleotides containing DNA adducts involves incorporating an appropriately modified base into an oligonucleotide and then modifying it with a chemical in order to generate the final desired product (Figure 2 - 3 (C)). Phosphoramidites containing post-synthetically displaceable leaving groups are commercially available, and researchers in the field can now, in many cases, easily gain access to their modified base of interest by simply displacing the leaving group in the oligonucleotide with an amine, alcohol, or thiol. This may be much more efficient than having to do a total synthesis of each phosphoramidite of interest. This approach usually works best with a relatively short oligonucleotide and a bulky adduct, in that separation of unreacted oligonucleotide will be necessary. This method was previously described in work published by the Manderville group, who performed Suzuki coupling of aryl boronic acid to an 8-Br-dG within an oligonucleotide to generate 8-aryl-dG modified oligonucleotides.

The last method used to synthesize oligonucleotides containing DNA adducts is the enzymatic approach (Figure 2 - 3 (D)). In this case, modified deoxynucleotide triphosphates (dNTPs) are prepared and incorporated into a primer strand by the action of a polymerase reading the template strand. Unfortunately, DNA polymerases often fail to properly incorporate modified dNTPs, and this method is not usually feasible on a large scale. The limited amount of oligonucleotides that can be produced may be sufficient for some biological studies, but the limitation may be an issue, even for characterization. The purification and analysis required for this method is not trivial.
2.2.1.1. Incorporation by the Phosphoramidite Approach

Nucleic acids are sensitive to a wide range of chemical reactions, and relatively mild reaction conditions are required for their chemical synthesis. This limits the range of chemical reactions one can employ in the synthesis to mild alkaline hydrolysis, very mild acidic hydrolysis, mild nucleophilic displacement reactions, base-catalysed elimination reactions and certain mild redox reactions.\textsuperscript{44}

The most convenient way to assemble an oligonucleotide is to utilise preformed deoxynucleoside phosphite derivatives as building blocks, and to couple these sequentially to a terminal nucleoside attached to a solid support. Since the 5′ hydroxyl group is a more effective nucleophile than the secondary 3′ hydroxyl group, the phosphite group is best placed on the 3′ position. Various strategies for the synthesis of modified nucleic acid phosphoramidites have been reported,\textsuperscript{36,45,46} with a general depiction of the natural nucleobase phosphoramidites shown in Figure 2 - 4.

![Phosphoramidite Structures](image)

Figure 2 - 4: General depiction of the unmodified nucleobase phosphoramidites. PG\textsubscript{2} used to depict a 5′OH protecting group while the nucleobase exocyclic amines are protected with a group symbolized by PG\textsubscript{1}.

When nucleosides are prepared for incorporation into oligonucleotides, it is usual to protect the nucleobase exocyclic amino groups first. This can be carried out with one of the following reagents: dimethylformamide diethyl acetal, isobutyric anhydride, 9-
fluorenylmethoxycarbonyl chloride (Fmoc-Cl), and 4-isopropylphenoxyacetyl chloride (iPrPac-Cl).

Once the nucleobases are protected, the 5'-hydroxyl is protected. This is typically carried out with the use of 4,4'-dimethoxytrityl chloride (DMT-Cl). This reagent is regioselective for the 5'-hydroxyl rather than the 3'-hydroxyl group due to steric effects and is readily removed under acidic conditions, liberating a bright orange-red DMT cation. The introduction of the phosphite group is the next step in the synthesis. Typically, this is performed using phosphite-triester chemistry and the phosphitylating agent 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite. Here, the 2-cyanoethyl group is the preferred protecting group because it can be removed conveniently and selectively at the end of the synthesis. Once the phosphoramidite building blocks are completely protected and synthesized, they are available for incorporation into oligonucleotides using a DNA synthesizer.

2.2.1.2. Automated DNA Synthesis

Oligonucleotide synthesis takes place in an automated process on a DNA synthesizer, a process which involves building the oligonucleotide on a solid support, one nucleotide at a time, in a repeated series of chemical reactions involving protection, coupling and de-protection steps. The first step of this process requires attachment of the 3' terminal deoxyribonucleoside of the oligonucleotide to a solid support. An example of this process is depicted in Figure 2 - 5, for which the protected 5'-O-DMT derivative was converted into its corresponding active succinate ester, and subsequently reacted with the amino groups on the support.
These solid supports are typically purchased with the desired starting nucleotide attached, unless the 3’ base is itself a modified base, in which it must be synthesized on a universal linker. Once attached to a solid support, the automated process may begin. The process involves four distinct steps: deprotection, condensation, capping and oxidation, and may be repeated as many times as necessary in order to yield an oligonucleotide of desired length (Scheme 2 - 4).
The first step, deprotection, hydrolyzes the 5'-O-DMT protecting group, allowing the 5'-OH to couple with an introduced phosphoramidite in the condensation/coupling stage. The condensation reaction is catalyzed by tetrazole, which protonates the N,N-diisopropyl phosphoramidite and converts the diisopropylamino moiety into a good leaving group. The protonated amino group is displaced by the 5' hydroxyl group of the support-bound mononucleotide and a dimer is formed.

In each coupling cycle around 1.5% of the oligonucleotide chains on the glass beads fail to react with the activated monomer and, if this situation were ignored, a complex mixture of
truncated sequences would accumulate. This is avoided by a capping step (Figure 8; step 3). Here, any unreacted 5′-hydroxyl groups are acetylated and thereby rendered inert to subsequent monomer additions.

After a capping step, the dimer is oxidized with aqueous iodine to convert the phosphite triester into a more stable phosphate triester (Step 4). The cycle is then continued by deprotecting the second base at the 5'OH by treatment with dichloroacetic acid (DCA) followed by the addition of the next desired base. This cycle is repeated the desired number of times, generating an oligonucleotide of desired length. Concentrated aqueous ammonia is then used to remove the protecting groups and cleave the oligonucleotide from the solid support, yielding fully deprotected oligonucleotide.

The average coupling efficiency in the phosphoramidite method of DNA synthesis, when carried out on an automatic machine, is approximately 98.5% and sequences up to around 100 bases can be prepared. Once the site-specifically modified oligonucleotides have been synthesized, they will be purified through a combination of several methods. Oligonucleotides are usually resolved based on charge using anion-exchange HPLC and an increasing ammonium acetate salt gradient, or based on hydrophobicity, using reversed-phase HPLC and an ion-pairing agent, such as triethylammonium acetate, with an increasing acetonitrile organic gradient.44

2.2.1.3. Interesting Modifications to Oligonucleotide Synthesis

Although oligonucleotides synthesized using typical phosphoramidite chemistry are of satisfactory quality for most biological uses, they are only marginally suitable for many recently discovered applications. These include large-scale synthesis, the preparation of oligonucleotides containing acid labile nucleobases, and the synthesis of genes or large gene fragments. The main issues are side reactions and coupling yields, which become more of an issue for longer sequences, and eventually result in unacceptable levels of truncated sequences to accumulate. Two of the most prevalent side reactions can be attributed to the acid exposure required for
detritylation.\textsuperscript{47} This step leads to depurination of protected bases and, because the detritylation step is reversible, a series of oligomers lacking one or more of the correct nucleotides.

Since its introduction, the DMT group has been widely used for the protection of 5’-OH functions in oligonucleotide synthesis. Its removal is carried out in di- and tri-chloroacetic acids, which cause depurination.\textsuperscript{47} Typically, depurination is a minor side reaction; however, it becomes more of an issue with purine derivatives that are prone to depurination. As discussed, 8-aryl derivatives of dG are known to be extremely prone to depurination under acidic conditions.\textsuperscript{34} Some of these modified nucleosides undergo depurination under acidic conditions up to 200 times more readily than dG,\textsuperscript{34} making assembly of their corresponding oligonucleotides extremely challenging. An optimal replacement of the DMT group must meet the following criteria: (i) it should be relatively easy to synthesize on large-scale from readily available (inexpensive) chemicals; (ii) it should offer the desired regioselectivity for the 5’-hydroxyl group in the nucleosides; (iii) it should rapidly cleave under mild acidic conditions, and (iv) it should offer crystalline products.

Researchers have proposed modifications to the standard solid phase synthesis cycle. One example involves a new two step synthesis cycle (Figure 2 - 6).\textsuperscript{48} This approach utilizes aqueous peroxy anions buffered under mildly basic conditions, pH 9.6 to remove an aryloxycarbonyl (ARCO) group, which is substituted for DMT in the four step cycle outlined earlier. Due to the strong nucleophilicity and mildly oxidizing nature of the peroxy anion solution, quantitative and irreversible removal of the 5’-O-carbonate protecting group and oxidation of the internucleotide phosphate triester can be carried out simultaneously. The authors also confirmed that oxidation of the heterocyclic nucleobases was not an issue. This two-step cycle uses fewer solvents, and eliminates acid catalyzed detritylation from the synthesis procedures. Consequently, depurination and reversibility of the detritylation step cease to exist as problems in oligomer synthesis. This method offers the advantage of never exposing the oligonucleotide to acid; however, the peroxy acid deprotection/oxidation step could present a
separate set of issues for C8-aryl-modified dG analogues, as they have been shown to have lower oxidation potentials than any of the natural nucleobases.

![Chemical structure diagram]

Figure 2 - 6: Solid-phase two step phosphoramidite oligodeoxynucleotide synthesis cycle.

Other possible replacements for the DMT group have been proposed.\textsuperscript{12} Reports within the literature established 2,7-dimethyl-9-phenyl-9H-xanthene (\textit{DMPx}) as an attractive alternative protecting group to the DMT protecting group.\textsuperscript{12,49,50}
The use of a pixyl (Px) group as a crystalline alternative to the DMT group for oligonucleotide synthesis was first proposed by Reese and Chattopadhyaya.\textsuperscript{47} Further investigations indicated that use of the Px group resulted in crystalline products, but at the cost of the formation of significant amounts of the undesired 3',5'\textsuperscript{-}bis-protected nucleoside.\textsuperscript{50} Px derivatives are more acid-sensitive than their DMT counterpart. Efforts to modulate the lability of the Px group resulted in the production of the 2,7-dimethylpixyl (DMPx) group. It was found that the DMPx group was more labile than the Px group during oligonucleotide synthesis.\textsuperscript{12}

Furthermore, Tram and coworkers conducted a study involving synthesizing a thymidine decamer (T\textsubscript{10}) with DMPx protected nucleosides.\textsuperscript{12} The optimization of the conditions required for oligonucleotide deprotection was based on this decamer, since thymidine is more difficult to detritylate than the other three deoxynucleosides.\textsuperscript{11} They tested various delivery times for
solutions of differing concentrations of di- and tri-chloroacetic acid. Their results demonstrated that acid-lability increases in the order of DMT, Px, and DMPx, as revealed by their half-life times of hydrolysis.12

Collectively, these reports suggested the DMPx protecting group as the most attractive alternative to DMT protection for the case of 8-aryl-dG adducts. This method seemed to be the most amenable of the modified approaches to the DNA synthesizer in our lab with the least amount of “new” chemistry being required. This prompted the current work, which entails the utilization of DMPx for protection of various 8-aryl-dG analogues in order to overcome their increased sensitivity to acid and increased rate of acid catalyzed depurination.

2.3. Materials and Methods

2.3.1. General Methods

A detailed outline of the general methods for experiments relating to Chapter 2 can be found in Appendix A.

2.3.2. Synthesis of C8-Aryl-dG Phosphoramidites

An outline for the synthesis of C8-aryl-dG phosphoramidites is provided in Scheme 2-5. Overall, the synthesis of the various adducts was conducted according to the literature,37 by palladium-catalyzed Suzuki-Miyaura cross-coupling reactions with 8-Br-dG and an appropriate boronic acid. The C8-aryl-dG adducts, 8-phenyl-2′-deoxyguanosine (PhdG), 8-(2-furyl)-2′-deoxyguanosine (FurG), 8-(8-quinolyl)-2′-deoxyguanosine (QdG), 8-(4″-cyanophenyl)-2′-deoxyguanosine (CNPhdG), 8-(2″-benzothienyl)-2′-deoxyguanosine (BthdG), and 8-pyrenyl-2′-deoxyguanosine (PydG), were then converted into their corresponding phosphoramidites as either their 5′-O-DMT or 5′-O-DMPx protected derivatives, as required.
8-Br-dG was synthesized according to the literature. To a 1 L Erlenmeyer flask was added dG (8.6782 g, 0.03042 mol), followed by the addition of 400 mL of 4:1 CH₃CN:H₂O (320 mL:80 mL). N-Bromosuccinimide (NBS) (7.7 g, 43 mmol) was slowly added to the mixture in roughly three equivalent portions, while stirring at room temperature. Each addition was made
after the yellow colour had almost completely faded from the reaction mixture. The reaction mixture was left to stir for 30 minutes at room temperature, and was then filtered under vacuum and washed with acetone. The filtered solid was then re-suspended in approximately 200 mL of acetone (in a 1 L Erlenmeyer flask), stirred for 2 h at room temperature and then cooled on ice. The resulting slurry was then filtered under vacuum and washed with small amounts of cold acetone, and 8-Br-dG was collected as a light pink solid (8.5925 g, 81.6 %). $^1$H (NMR) (DMSO-$d_6$) (300 MHz), δ: 10.80 (s, 1H), 6.49 (s, 2H), 6.16 (t, $J = 7.4$ Hz, 1H), 5.25 (d, $J = 4.5$ Hz, 1H), 4.86 (t, $J = 6.0$ Hz, 1H), 4.40 (bs, 1H), 3.80 (bs, 1H), 3.61 (m, 1H), 3.52 (m, 1H), 3.17 (m, 1H), 2.12 (m, 1H); $^{13}$C NMR (DMSO-$d_6$) (75.5 MHz), δ: 155.4, 153.3, 152.0, 120.6, 117.5, 87.9, 85.1, 71.0, 62.1, 36.5. Spectra obtained matched the published $^1$H NMR and $^{13}$C NMR data.$^{51}$

**Quinoline-8-Boronic acid**

Quinoline-8-boronic acid was also synthesized according to literature procedures$^{52}$ to avoid the costly purchase from Frontier Scientific or Sigma Aldrich.

8-Bromo quinoline (2.68 g, 12.8 mmol) in dry THF (10 mL) was cooled to -78°C under argon and $^9$BuLi (1.55 M in hexane, 10 mL, 15.5 mmol) was added dropwise over 30 min. After 1 hour trimethyl borate (3 mL, 26.9 mmol) was added dropwise and the ice bath was removed. The mixture was stirred for 1 hour at room temperature and then 3 M HCl (30 mL) was added. The aqueous layer was washed with Et$_2$O and neutralized with solid NaHCO$_3$. The resulting brown precipitate was collected and recrystallized from acetone/hexanes to give product as a pale yellow solid in 60% yield. MP: 67-70°C. $^1$H NMR (300 MHz, d$_6$-CD$_3$OD) δ =9.29 (dd, $J = 1.7$, 5.3 Hz, 1H), 8.52 (dd, $J = 1.7$, 8.3 Hz, 1H), 8.10 (dd, $J = 1.5$, 6.8 Hz, 1H), 7.80 (dd, $J = 1.4$, 8.2 Hz, 1H), 7.71 (dd, $J = 5.3$, 8.3 Hz, 1H), 7.60 (dd, $J = 6.8$, 8.2 Hz, 1H); $^{13}$C NMR 146.3, 144.7, 143.3, 134.9, 130.6, 129.4, 128.0, 121.6. HRMS calc’d for C$_9$H$_8$BNO$_2^+$ [M+H$^+$] 174.0726; found 174.0724.
2.3.2.1. Suzuki-Miyaura Coupling of 8-Br-dG with Boronic Acids

2.3.2.1.1. General Procedure

These reactions were conducted according to the literature, and are briefly described here. Pd(OAc)$_2$ (22 mg, 0.1 mmol), TPPTS (148 mg, 25 mmol), Na$_2$CO$_3$ (800 mg, 7.5 mmol), 8-bromo-dG (1.30 g, 3.75 mmol) and aryl boronic acid (11.25 mmol) were placed in a round-bottomed flask fitted with a condenser and reverse-filled with argon. Degassed 2:1 H$_2$O:CH$_3$CN (35 mL) solution was added and the solution was heated to 100°C for 4-8 hours and monitored with TLC. Following completion, the mixture was diluted with 200 mL of H$_2$O and pH was adjusted to 7.5 with 1.0 M aqueous HCl. The mixture was then cooled to 0°C, filtered and dried to yield product.

8-(Furan-2-yl)-2'-deoxyguanosine (FurG) (1a):

Synthesis was completed as described to yield 1.04 g (83%) of a white grey powder. MP = 223-225°C (d). $^1$H NMR (DMSO-d$_6$) (300 MHz) δ = 10.75 (bs, 1H), 7.89 (s, 1H), 6.94 (d, $J = 3.3$ Hz, 1H), 6.78 (m, 1H), 6.42 (s, 2H), 6.35 (t, $J = 7.4$ Hz, 1H), 5.18 (d, $J = 4.2$ Hz, 1H), 4.94 (t, $J = 5.7$ Hz, 1H), 4.37 (bs, 1H), 3.79 (bs, 1H), 3.58 (m, 1H), 3.48 (m, 1H), 3.15 (m, 1H), 2.06 (m, 1H); $^{13}$C NMR (DMSO-d$_6$) (300 MHz) δ = 156.4, 153.1, 151.6, 144.4, 144.1, 137.8, 117.4, 112.0, 111.8, 87.9, 84.4, 71.1, 62.1, 37.0; HRMS calcd for C$_{14}$H$_{15}$N$_3$O$_5$ [M+H$^+$] 334.1151, found 334.1142.

8-(Phenyl)-2'-deoxyguanosine (PhG) (1b):

Synthesis was completed as described to yield 1.03 g (80% yield) of a white powder. MP = 204-206°C. $^1$H NMR (300 MHz, DMSO-d$_6$) δ = 10.78 (s, 1H), 7.62-7.51 (m, 5H), 6.40 (s, 2H), 6.05 (t, $J = 9$ Hz, 1H), 5.11 (d, $J = 6$Hz, 1H), 4.98 (t, $J = 6$Hz, 1H), 4.32 (m, 1H), 3.77 (m, 1H), 3.62 (m, 1H), 3.54 (m, 1H) 3.13 (m, 1H), 1.99 (m, 1H); $^{13}$C NMR (600 MHz, DMSO-d$_6$) δ =
156.8, 153.1, 151.9, 147.0, 130.3, 129.4, 129.1, 128.6, 117.1, 87.9, 84.6, 71.2, 62.1, 36.5. HRMS calcd C₁₆H₁₈N₅O₄⁺ [M+H⁺] 344.1353; found 344.1349. NMR spectra matched literature.³⁶

8-(4''CN-Phenyl)-2'-dG (CNpG) (1c):

Synthesis was completed as described with a modification to the solvent used. Instead of using 35 mL 2:1 H₂O:CH₃CN the reaction was optimized with a 1:1:1 H₂O:CH₃CN:MeOH solvent. Upon completion of the reaction this mixture was partially rotovapped to remove as much methanol as possible prior to the pH adjustment. The reaction yielded 1.21 g (88%) of a yellowish-white powder MP = 208-210°C. ¹H NMR (300MHz, DMSO-d₆) δ = 7.97 (d, J = 8.3 Hz, 2H), 7.87 (d, J = 8.3 Hz, 2H), 6.87 (bs, 2H), 6.08 (t, J = 7.4Hz, 1H), 5.52 (bs, 1H), 5.18 (bs, 1H), 4.36 (m, 1H), 3.82 (m, 1H), 3.68 (dd, J = 4.3, 12 Hz, 1H), 3.53 (dd, J = 4.3, 12 Hz, 1H), 3.16 (m, 1H), 2.05 (m, 1H). ¹³C NMR (300MHz, DMSO-d₆) δ = 160.7, 156.0, 152.6, 144.4, 135.0, 133, 132.5, 129.6, 128.0, 118.6, 117.9, 111.4, 88.1, 84.8, 71.2, 62.2, 36.9. HRMS calcd for C₁₇H₁₇N₆O₄⁺ [M+H⁺] 369.1306, found 369.1321.

8-(8''-Quinolyl)-2'-dG (QdG) 1d:

Similar to 1c, the synthesis of 1d was optimized with a 1:1:2 H₂O:CH₃CN:MeOH solvent. Upon completion of the reaction this mixture was partially rotovapped to remove as much methanol as possible prior to the pH adjustment. The reaction yielded 2.2g (73%) of a yellowish-white powder. MP = 195-198°C (d). ¹H NMR (600 MHz, DMSO-d₆) δ = 10.74 (s, 1H), 8.89 (d, J = 2.6Hz, 1H), 8.50 (dd, J = 1.4, 8.3 Hz, 1H), 8.19 (dd, J = 1, 8.3 Hz, 1H), 7.92 (bs, 1H), 7.74 (t, J = 7.7Hz, 1H), 7.61 (dd, J = 4.1, 8.3Hz, 1H), 6.31 (bs, 2H), 5.51 (bs, 1H), 5.04 (bs, 1H), 4.90 (d, J = 1.1Hz, 1H), 4.27 (bs, 1H), 3.55 (m, 2H), 3.41 (bs, 1H), 3.18 (bs, 1H), 2.09 (bs, 1H). ¹³C NMR (600 MHz, DMSO-d₆) δ = 157.7, 157.2, 151.3, 150.5, 146.9, 146.3, 136.9, 132.8, 130.9, 129.7, 128, 126.6, 122.5, 120.8, 87.8, 71.2, 62.4, 38.0. HRMS calcd for C₁₉H₁₉N₆O₄⁺ [M + H⁺] 395.1462; found 395.1468.
8-(2''-Benzothienyl) 2'-deoxyguanosine (BthdG) 1e:

Synthesis was completed as described to yield 1.03 g (80% yield) of a white powder. MP = 220-224 (d) °C. $^1$H NMR (400Mhz, DMSO-d$_6$) δ = 10.99 (bs, 1H), 8.02-7.99 (m, 1H), 7.95-7.86 (m, 2H), 7.44-7.38 (m, 2H), 6.62 (bs, 2H) 6.42 (t, J = 7.24 Hz, 1H), 5.23 (bs, 1H), 5.06 (bs, 1H), 4.43 (bs, 1H), 3.85 (q, J = 4.7 Hz, 1H), 3.71 (m, 1H), 3.60 (m, 1H), 3.27-3.20 (m, 1H), 2.14-2.08 (m, 1H); $^{13}$CNMR (400 MHz, DMSO-d$_6$) δ = 156.8, 153.6, 152.6, 140.9, 139.7, 139.3, 132.5, 125.5, 124.9, 124.5, 124.4, 122.3, 117.3, 87.8, 84.2, 70.7, 61.6, 36.7. HRMS calcd for C$_{18}$H$_{18}$N$_5$O$_4$S$^+$ [M+H$^+$]: 400.1074; found 400.1096.

8-(Pyrenyl)-2'-deoxyguanosine (PydG) 1f:

Similar to 1c, the synthesis was completed as described with a modification to the solvent used. Instead of using 35 mL 2:1 H$_2$O:CH$_3$CN, the reaction was optimized with a 1:1:2 H$_2$O:CH$_3$CN:MeOH solvent. Upon completion of the reaction, this solvent was partially rotovapped to remove as much methanol as possible prior to the pH adjustment. Material was successively washed with hexane, Et$_2$O, EtOAc, MeOH and acetone to remove unreacted boronic acid. The reaction yielded 60% of an off white powder. MP = 247-250°C (d). $^1$H NMR (600 MHz, DMSO-d$_6$) δ = 11.60 (bs, 1H), 8.44 (d, J = 7.7 Hz, 1H), 8.39 (d, J = 7.6 Hz, 1H), 8.37 (d, J = 7.5 Hz, 1H), 8.33 (d, J = 9 Hz, 1H), 8.30 (d, J = 9.0 Hz, 1H), 8.26 (d, J = 9.2Hz, 1H), 8.18-8.13 (m, 2H), 8.02 (d, J = 9.2Hz, 1H), 5.80 (m, 1H), 5.16-5.02 (m, 2H), 4.25 (bs, 1H), 3.64 (m, 1H), 3.55 (m, 1H), 3.47 (m, 1H), 1.99 (m, 1H); $^{13}$C NMR (600 MHz, DMSO-d$_6$) δ = 158.0, 155.1, 150.7, 146.4, 131.8, 130.9, 130.8, 130.4 130.2 128.8, 128.7, 127.5, 127.4, 127.0, 126.4, 126.2, 126.0, 125.2, 124.8, 124.7, 123.9, 123.7, 121.0, 88.0, 85.3, 71.2, 62.3, 37.8. HRMS calcd for C$_{26}$H$_{22}$N$_5$O$_4$$^+$ [M+H$^+$] 468.1666; found 468.1681.
2.3.2.2. \(N^2\) Protection.

Past research within the Manderville group had utilized a transient protection method employing TMS-Cl and isobutyric anhydride for \(N^2\) protection. The choice for utilizing formamidine protection for the \(N^2\) position stemmed from findings that this protection strategy increased the resistance of purines to depurination upon protection.\(^{53}\) Replacement of MeOH with DMF as a solvent decreased the reaction time from 5 days to 5 hours, and product was able to be used without further purification as the quantitative reactions were routinely achieved.

2.3.2.2.1. General Procedure

A typical procedure involved placing 8-aryl-2′-deoxyguanosine (3.3 mmol) in a round bottom flask (RBF) and reverse-filled with argon. Reactions were found to progress fastest in polar aprotic solvents such as dry DMF as opposed to polar protic solvents such as MeOH, as had been reported within the literature.\(^{11,53}\) 15 mL of dry DMF was added, followed by dimethylformamide-diethyl-acetal (2.7 mL, 13.5 mmol), and the mixture was allowed to go to completion with stirring at room temperature. The reaction mixture was then evaporated to dryness and the solid washed with MeOH and dried to yield product which was converted quantitatively and usable without further purification.

\(N^2\)-(Dimethylformamidyl)-8-(2″-Furyl)-2′-dG (2a).

Synthesis performed as outlined afforded 2a as a gray solid (1.28 g, 99.0% yield). MP = 195–197 °C. \(^1\)H NMR (600 MHz, DMSO-\(d_6\)) \(\delta = 11.59\) (s, 1H), 8.57 (s, 1H), 8.01 (d, \(J = 1.2\) Hz, 1H), 7.07 (d, \(J = 3.3\) Hz, 1H), 6.78 (dd, \(J = 1.8\) Hz, 3.4 Hz, 1H), 6.51 (t, \(J = 7.2\) Hz, 1H), 6.36 (d, \(J = 4.3\) Hz, 1H), 4.95 (t, \(J = 5.1\) Hz, 1H), 4.54 (m, 1H), 3.98 (m, 1H), 3.88 (m, 1H), 3.71 (m, 1H), 3.27 (m, 1H), 3.22 (s, 3H), 3.11 (s, 3H), 2.23 (m, 1H); \(^13\)C NMR (600 MHz, DMSO-\(d_6\)) \(\delta = 159.2, 158.3, 158.0, 151.3, 145.6, 145.0, 139.8, 121.4, 113.3, 112.8, 88.6, 85.5, 71.9, 62.9, 41.8, 38.4, 35.6. HRMS calcd for \(C_{27}H_{30}N_6O_5^+\) [M + H+] 389.1573; found 389.1568.
$N^2$-(Dimethylformamidyl)-8-Phenyl-2'-dG(2b)

Synthesis performed as outlined afforded 1.25 g (3.14 mmol) of a white powder corresponding to a 95% yield. MP = 218–220°C. $^1$H NMR (300 MHz, DMSO-d$_6$) $\delta$ = 11.46 (s, 1H), 8.49 (s, 1H), 7.65-7.53 (m, 5H), 6.09 (t, $J$ = 6 Hz, 1H), 5.21 (d, $J$ = 6.0 Hz, 1H), 4.88 (t, $J$ = 6.0 Hz, 1H), 4.42 (bs, 1H), 3.81 (bs, 1H), 3.67 (m, 1H), 3.61 (m, 1H) 3.33 (m, 1H), 3.22 (s, 3H), 3.04 (s, 3H), 2.08 (s, 1H); $^{13}$C NMR (300 MHz, DMSO-d$_6$) $\delta$ = 158.1, 157.5, 156.8, 150.7, 148.1, 130.1, 129.5, 129.1, 128.7, 120.1, 87.7, 84.8, 71.0, 62.0, 40.8, 37.0, 34.6. Spectra matched that of the published literature.$^{36}$

$N^2$-(Dimethylformamidyl)-8-(4''-Cyanophenyl)-2'-dG (2c).

Isolated 1.33g (3.14 mmol) of a yellow powder corresponding to a 95% yield. MP 179–182°C. $^1$H NMR (600 MHz, DMSO-d$_6$) $\delta$ = 8.48 (s, 1H), 8.00 (d, $J$ = 8.3 Hz, 2H), 7.85 (d, $J$ = 8.3 Hz, 2H), 6.11 (t, $J$ = 7.7 Hz, 1H), 5.75 (bs, 1H), 5.30 (bs, 1H), 4.44 (m, 1H), 3.86 (m, 1H), 3.67 (dd, $J$ = 3.5, 12 Hz, 1H), 3.55 (dd, $J$ = 3.8, 11.9 Hz, 1H), 3.24 (m, 1H), 3.09 (s, 3H), 3.00 (s, 3H), 2.07 (m, 1H); $^{13}$C NMR (600 MHz, DMSO-d$_6$) $\delta$ = 157.7, 151.6, 145.0, 142.5, 141.9, 135.0, 133.0, 132.6, 129.7, 128.1, 121.3, 118.6, 111.4, 88.3, 85.2, 71.3, 62.3, 40.5, 37.3, 34.4. HRMS calcd for C$_{20}$H$_{22}$N$_7$O$_4$ [M+H$^+$] 424.1733, found 424.1724.

$N^2$-(Dimethylformamidyl)-8-(8''-Quinolyl)-2'-dG (2d).

Synthesis performed as outlined afforded 1.39g (3.1 mmol) as a yellow powder corresponding to a 94% yield. MP = 188–192°C. $^1$H NMR (300 MHz, DMSO-d$_6$) $\delta$ = 11.45 (bs, 1H), 8.90 (m, 1H), 8.50 (m, 2H), 8.21 (d, $J$ = 8.2 Hz, 1H), 7.95 (d, $J$ = 6.7 Hz, 1H), 7.75 (t, $J$ = 7.7 Hz, 1H), 7.62 (dd, $J$ = 4.2, 8.2 Hz, 1H), 5.65 (bs, 1H), 4.97 (m, 1H), 4.89 (m, 1H), 4.30 (bs, 1H), 4.09 (m, 1H), 3.57-3.45 (m, 2H), 3.16-3.12 (m, 4H), 3.03 (s, 3H), 2.10 (bs, 1H); $^{13}$C NMR (300 MHz, DMSO-d$_6$) $\delta$ = 158.1, 157.4, 157.0, 151.1, 150.2, 146.7, 146.1, 136.6, 132.6, 130.5, 129.4, 127.7, 126.3, 122.0, 120.6, 87.5, 85.4, 70.9, 62.1, 48.6, 37.4, 34.5. HRMS calcd for C$_{22}$H$_{23}$N$_7$O$_4$K$^+$ [M+K$^+$] 488.1449; found 488.1438.
$N^2$-(Dimethylformamidyl)-8-(2′′-Benzothienyl)-2′-dG (2e):

Synthesis performed as outlined afforded 1.45 g of a beige powder (97% yield). MP = 213-218°C. $^1$H NMR (400MHz, DMSO-d$_6$) δ = 11.57 (bs, 1H), 8.54 (s, 1H), 8.04-8.01 (m, 1H), 7.97-7.90 (m, 2H), 7.43 (m, 2H), 6.49 (t, J = 7.2 Hz, 1H), 5.33 (bs, 1H), 5.01 (bs, 1H), 4.53 (m, 1H), 3.86 (q, J = 4.4 Hz, 1H), 3.71 (m, 1H), 3.60 (m, 1H), 3.28 (m, 1H), 3.16 (s, 3H), 3.05 (s, 3H), 2.21-2.15 (m, 1H). $^{13}$C NMR (400 MHz, DMSO-d$_6$): 158.2, 157.6, 157.3, 151.3, 141.9, 139.6, 139.4, 132.3, 125.6, 124.9, 124.6, 124.5, 122.3, 120.4, 87.7, 84.4, 70.6, 61.6, 40.9, 37.2, 35.8. HRMS Calcd for C$_{21}$H$_{23}$N$_6$O$_4$S$^+$ [M+H$^+$]: 455.1496; found 455.1503.

$N^2$-(Dimethylformamidyl)-8-(Pyrenyl)-2′-dG (2f):

Synthesis performed as outlined afforded 1.62 g of an off white powder (94% yield). MP = 207-210°C. $^1$H NMR (600MHz, DMSO-d$_6$) δ = 11.58 (bs, 1H), 8.54 (s, 1H), 8.44 (d, J = 7.9 Hz, 1H), 8.40 (d, J = 7.6Hz, 1H), 8.36 (d, J = 7.5 Hz, 1H), 8.32 (d, J = 9 Hz, 1H), 8.28 (d, J = 9 Hz, 1H), 8.25 (d, J = 9.3 Hz, 1H), 8.18-8.13 (m, 2H), 8.02 (d, J = 9.2 Hz, 1H), 5.80 (m, 1H), 5.17 (bs, 1H), 5.02 (bs, 1H), 4.25 (bs, 1H), 3.63 (m, 1H), 3.55 (m, 1H), 3.47 (m, 1H), 3.14 (s, 3H), 3.04 (s, 3H), 1.99 (m, 1H). $^{13}$C NMR (600Mhz, DMSO-d6) δ = 158.1, 150.6, 131.7, 130.8, 130.6, 130.3, 130.1, 128.7, 128.6, 128.5, 127.4, 127.3, 126.8, 126.3, 126.1, 125.9, 125.1, 124.7, 124.6, 123.8, 123.5, 120.8, 87.8, 85.2, 71.1, 62.1, 40.7, 40.1, 37.6, 35.8, 34.6. HRMS calcd for C$_{29}$H$_{27}$N$_6$O$_4$S$^+$ [M+H$^+$]: 523.2088; found 523.2098.

2.3.2.3. 5′OH DMT Protection

DMT protection was slightly modified from protocols reported within the literature. Selective protection of the 5′OH was more difficult to achieve with 8-aryl-dG adducts than with unmodified nucleosides, presumably as a result of added steric bulk conferred by the aryl ring and differences in solubility. This enabled the 3′OH to compete kinetically with the 5′OH for the reaction with DMT-Cl, and thus resulted in low yields. The reaction also yielded, to some extent, some 3′OH and 5′OH doubly protected material, which was isolated but was rendered unusable.
It was found that premixing of the DMT-Cl with pyridine in a separate vessel and increasing the delivery time to at least 30 minutes increased the overall yield. In some cases the nucleoside was dissolved in DMF rather than pyridine for solubility considerations. Carrying out the reaction at 0°C rather than room temperature also resulted in marginal increases in yield. It was also found that the addition of TEA to the crude material prior to rotary evaporation and to the solvent used as the mobile phase in column chromatography was required to prevent deglycosylation upon concentration and purification with silica.

2.3.2.3.1. General Procedure

N²-(Dimethylformamidyl)-8-aryl-2’-deoxyguanosine (2.7 mmol) was co-evaporated from dry pyridine (3 x 5 mL) in a RBF. The RBF was then fitted with a constant pressure dropping funnel, reverse filled with argon, and 7 mL of dry pyridine (or DMF) was added to the RBF and cooled to 0°C. A DMT-Cl (1.28 g, 3.78 mmol) pyridine (5 mL) solution was added to the dropping funnel under argon and allowed to add dropwise over 30 min. The reaction was allowed to stir at room temperature under argon and was monitored by TLC. Upon completion, the mixture was diluted with methylene chloride (10 mL) and washed with water (2 x 10 mL). TEA (1 mL) was added and the mixture was evaporated to yield an oil. The oil was then loaded onto a silica column and run with 95:5 CH₂Cl₂:TEA to elute unreacted DMT material; product was then eluted with MeOH:CH₂Cl₂:TEA (5:90:5), fractions deemed pure were combined and rotovapped to yield a powder.

5’-O-(4,4’-Dimethoxytrityl)-N²-(Dimethylformamidyl)-8-(2”-Furyl)-2’-dG(3a):

Synthesis performed as outlined afforded product as a white solid (1.45 g, 78.0% yield):
MP = 162–167 °C; ¹H NMR (600.1 MHz, CD₃CN) δ = 9.44 (bs, 1H), 8.33 (s, 1H), 7.66 (s, 1H), 7.29 (m, 2H), 7.16 (m, 7H), 7.06 (d, J = 3.4 Hz, 1H), 6.7 (m, 4H), 6.62 (dd, J = 1.62 Hz, 3.24 Hz,1H), 6.50 (dd, J = 4.5 Hz, 8 Hz, 1H), 4.77 (dd, J = 5.9 Hz, 13.0 Hz, 1H), 3.97 (m, 1H), 3.71 (s, 3H), 3.70 (s, 3H), 3.46 (m, 1H) 3.25 (m, 1H), 3.14 (m, 1H), 3.14 (s, 3H), 3.02 (s, 3H), 2.24 (m,
1H). $^{13}$C NMR (151 MHz, CD$_3$CN) $\delta$ = 160.0, 159.9, 159.4, 158.9, 157.9, 152.1, 146.7, 146.3, 145.6, 141.5, 137.6, 137.4, 131.4, 131.1, 129.4, 129.1, 128.1, 122.2, 118.8 (CD$_3$CN), 114.3, 113.7, 113.2, 87.5, 87.1, 86.1, 72.9, 65.6, 56.32, 56.3, 42.2, 38.9, 35.7, 1.8 (CD$_3$CN); HRMS calcd for C$_{38}$H$_{38}$N$_6$O$_7$ [M + H$^+$] 691.2880, found 691.2866.

$5'$-O-(4,4'-Dimethoxytrityl)-N$^2$-(Dimethylformamidyl)-8-Phenyl-2'-dG (3b):

Synthesis performed as outlined afforded product as a white powder (1.32 g, 70% yield). MP = 196-200°C. $^1$H NMR (600 MHz, DMSO-d$_6$) $\delta = 11.44$ (s, 1H), 8.28 (s, 1H), 7.75 (m, 2H), 7.51 (m, 3H), 7.30 (m, 2H), 7.16 (m, 7H), 6.74 (m, 4H), 6.14 (t, J = 6 Hz, 1H), 5.27 (d, J = 6 Hz, 1H), 4.56 (s, 1H), 3.89 (m, 1H), 3.69 (s, 3H), 3.68 (s, 3H), 3.27 (m, 1H), 3.19 (m, 1H), 3.11 (m, 1H), 3.01 (s, 3H), 2.96 (s, 3H), 2.18 (m, 1H). $^{13}$C NMR (600 MHz, DMSO-d$_6$) $\delta$ = 157.9, 157.8, 157.6, 157.5, 156.3, 150.4, 148.2, 144.9, 135.7, 135.6, 130.3, 129.6, 129.5, 129.4, 129.2, 128.6, 127.6, 126.5, 120.1, 112.9, 112.9, 85.8, 85.1, 84.1, 70.9, 63.9, 54.9, 54.9, 45.6, 40.8, 40.0, 37.4, 34.6. HRMS calcd for C$_{40}$H$_{41}$N$_6$O$_6$ $^+$ [M+H$^+$] 701.3082; found 701.3078. Spectra matched that of the literature.

$5'$-O-(4,4'-Dimethoxytrityl)-N$^2$-(Dimethylformamidyl)-8-(4''-CN-Phenyl)-2'-dG(3c):

Synthesis performed as outlined afforded product as a white powder (1.42 g, 73% yield). MP = 166-170°C. $^1$H NMR (300 MHz, DMSO-d$_6$) $\delta = 8.36$ (s, 1H), 8.01-7.97 (d, J = 8.6 Hz, 2H), 7.72-7.69 (d, J = 8.6 Hz, 2H), 7.42-7.39 (m, 2H), 7.31-7.19 (m, 7H), 6.79-6.74 (m, 4H), 6.24-6.19 (m, 1H), 4.69 (m, 1H), 4.07 (m, 1H), 3.76 (s, 3H), 3.75 (s, 3H), 3.50 (m, 1H), 3.36-3.27 (m, 2H), 3.05 (s, 3H), 2.94 (s, 3H), 2.22 (m, 1H). $^{13}$C NMR (300 MHz, DMSO-d$_6$) $\delta$ = 159.0, 158.4, 158.4, 156.8, 152.1, 147.9, 145.4, 136.3, 136.2, 135.1, 132.6, 130.4, 130.3, 128.5, 128.2, 127.2, 121.3, 118.8, 113.4, 113.3, 86.6, 86.2, 84.6, 72.6, 64.8, 55.6, 41.6, 38.2, 35.3. HRMS calcd for C$_{41}$H$_{40}$N$_7$O$_6$ $^+$ [M+H$^+$] 726.3040, found 726.3060.
5′-O-(4,4′-Dimethoxytrityl)-N2-(Dimethylformamidyl)-8-(8″-Quinoly)-2′-dG (3d).

Synthesis performed as outlined afforded 1.12 g of a yellow powder corresponding to a 55% yield. MP = 178–181°C. 1H NMR (600 MHz, DMSO-d6) δ = 11.39 (bs, 1H), 8.88 (bs, 1H), 8.50 (d, J = 7.8 Hz, 1H), 8.19 (m, 2H), 8.02 (bs, 1H), 7.78 (bs, 1H), 7.62 (m, 1H), 7.26 (m, 2H), 7.17-7.11 (m, 7H), 6.76 (m, 4H), 5.65 (bs, 1H), 5.08 (bs, 1H), 4.42 (bs, 1H), 3.70 (s, 6H), 3.61 (m, 1H), 3.21 (bs, 1H), 3.07 (m, 1H), 3.00 (m, 4H), 3.91 (s, 3H), 1.78 (bs, 1H); 13C (600MHz, DMSO-d6) δ = 158.0, 157.9, 157.6, 151.2, 150.2, 145.0, 136.7, 135.7, 135.6, 132.8, 130.5, 129.7, 129.4, 127.8, 127.6, 126.5, 122.1, 113.0, 112.95, 87.0, 86.5, 85.2, 71, 64.5, 55.0, 54.97, 45.7, 40.7, 40.1, 37.9, 34.6, 30.7. HRMS calcd for C43H42N7O6+ [M+H]+: 752.3197; found 752.3191.

5′-O-(4,4′-Dimethoxytrityl)-N2-(Dimethylformamidyl)-8-(2″-Benzothienyl)-2′-dG (3e).

Synthesis performed as outlined afforded product as an off white powder (1.40 g, 70%). MP = 185-190°C (d). 1H NMR (400MHz, CDCl3) δ = 9.13 (bs, 1H), 8.35 (s, 1H), 7.84 (s, 1H), 7.78 (m, 1H), 7.67 (m, 1H), 7.35-7.30 (m, 4H), 7.24-7.20 (m, 4H), 7.17-7.11 (m, 3H), 6.70-6.67 (m, 4H), 6.43 (dd, J = 7.9, 4.6 Hz, 1H), 4.94 (q, J = 6.2Hz), 4.11 (q, J = 5.5 Hz, 1H), 3.68 (s, 6H), 3.52-3.43 (m, 2H), 3.35 (m, 1H), 2.95 (s, 3H), 2.89 (s, 3H), 2.32 (m, 1H). 13C NMR (400MHz, CDCl3) δ = 158.3, 157.8, 157.5, 155.8, 151.1, 144.7, 143.9, 140.5, 139.6, 135.7, 135.67, 132.2, 129.9, 129.8, 128.0, 127.7, 126.7, 125.3, 125.2, 124.4, 124.3, 122.1,120.8, 112.9, 86.1, 85.2, 84.0, 72.7, 64.2, 55.1, 41.3, 37.6, 35.0. HRMS calcd for C42H41N6O6S+ [M+H]+: 757.2803; found 757.2821.

5′-O-(4,4′-Dimethoxytrityl)-N2-(Dimethylformamidyl)-8-(Pyrenyl)-2′-dG (3f): Synthesis performed as outlined afforded product as an off white powder (1.11g, 50%). MP = 198-202°C (d). 1H NMR (600 MHz, CDCl3) δ = 8.31 (bs, 1H), 8.17-7.92 (m, 9H), 7.31 (d, J = 7.4 Hz, 2H), 7.18-7.13 (m, 7H), 6.66 (m, 4H), 6.02 (bs, 1H), 3.73 (m, 1H), 3.68 (s, 3H), 3.67 (s, 3H), 3.31 (m, 1H), 3.23 (m, 1H), 3.01 (m, 1H), 2.93 (m, 3H), 2.82 (bs, 3H), 2.02 (m, 1H). 13C
NMR (600MHz, CDCl$_3$) $\delta = 158.6, 158.3, 158.0, 156.2, 151.1, 148.2, 145.0, 136.1, 132.3, 131.3, 130.9, 130.8, 130.1, 129.0, 128.88, 128.7, 128.3, 127.9, 127.4, 126.9, 126.4, 125.9, 125.8, 125.2, 124.6, 124.4, 124.2, 120.8, 113.2, 86.3, 85.0, 72.2, 64.6, 55.4, 41.4, 38.6, 35.2. HRMS calcd for C$_{50}$H$_{45}$N$_6$O$_6^+$ [M+H$^+$]: 825.3395; found 825.3404.

2.3.2.4. Alternative 5′OH Protection with DMPx

Before nucleosides could be protected with DMPx, DMPx-Cl needed to be synthesized, as it is not commercially available. The general route for the synthesis of DMPx-Cl is shown in Scheme 2 - 6 and follows the literature. Key starting materials were purchased from Sigma Aldrich. Under the optimized conditions, reaction of di-$p$-tolyl ether and benzoic acid in the presence of zinc chloride and phosphorus oxychloride at 95°C furnished 2,7-dimethyl-9-phenylxanthen-9-methyl ether (DMPx-OCH$_3$) in 60% yield. Treatment of DMPx-OCH$_3$ with excess acetyl chloride in toluene furnished DMPx-Cl in quantitative yield.

Scheme 2 - 6: Synthesis of DMPx – Cl
It is also important to note that the corresponding unmodified dG, dC, dA and dT nucleoside phosphoramidites (Figure 2-9) needed to be synthesized with DMPx protection for use in the modified DNA synthesis protocols. Unmodified DMF-dG, Ac-dC, DMF-dA and T nucleosides were purchased and protected with the 5'-O-DMPx protocol described below before conversion to their corresponding phosphoramidite.

Figure 2-9: Synthesis of DMPx protected nucleosides for utilization in solid phase synthesis.
Synthesis of DMPx-Cl

Synthesis was performed according to the literature.\textsuperscript{12} \textit{p}-Tolyl ether (5 g, 0.025 mol), benzoic acid (3.75 g, 0.03 mol), zinc chloride (6 g, 0.75 mol) and phosphorous oxychloride (7.5 mL, 0.08 mol) were heated at 95°C for 3 h. The mixture was then cooled to room temperature and ethyl acetate (10 mL) was added to form a suspension. The suspension was poured into 250 mL stirring DI water at room temperature. The mixture was heated under reflux for 15 minutes and cooled to room temperature overnight. It was then filtered and washed with water (200 mL). The damp cake was suspended with 100 mL methanol and stirred while heating to boil for 2 or 3 minutes. The resultant suspension was allowed to cool to room temperature over a period of 3 hours and was then filtered, washed with methanol and dried to give the compound as a white solid, which was found to be DMPx-OMe in 64% yield. MP = 135–137°C. \textsuperscript{1}H NMR (300 MHz, CDCl\textsubscript{3}) $\delta$ = 7.57-7.53 (m, 2H), 7.45-7.40 (m, 2H), 7.35-7.33 (m, 1H), 7.24 (m, 4H), 7.14 (s, 2H), 3.09 (s, 3H), 2.39 (s, 6H). \textsuperscript{13}C NMR (300 MHz, CDCl\textsubscript{3}) $\delta$ = 149.9, 149.3, 132.7, 130.1, 129.9, 129.4, 129.3, 128.0, 126.6, 126.57, 122.7, 116.1, 76.3, 51.2, 21.0. DMPx-OMe (2.0 g, 6.9 mmol) was then azeotroped with dry toluene (2 x 10 mL) and the residue was dissolved in dry toluene (10 mL). Acetyl chloride (5.8 mL, 81.8 mmol) was added and the reaction was left to stir for 16 h at room temperature under argon. Upon completion the reaction was evaporated to dryness and co-evaporated with dry toluene (3 x 10 mL). The product was further dried under vacuum at 40°C for 3 hours to afford DMPx-Cl in quantitative yield as an orange/brown solid. MP 108-109°C.\textsuperscript{1}H NMR (300 MHz, CD\textsubscript{2}Cl\textsubscript{2}) $\delta$ = 7.63-7.60 (m, 2H), 7.45-7.42 (m, 2H), 7.29-7.25 (m, 4H), 7.05 (s, 2H), 2.26 (s, 6H). \textsuperscript{13}C NMR (300 MHz, CDCl\textsubscript{3}) $\delta$ = 149.3, 143.5, 134.4, 133.1, 130.4, 129.2, 128.5, 125.3, 117.1, 85.7, 21.0. Spectra were consistent with those of literature reports.\textsuperscript{12}
2.3.2.4.1. General Procedure for DMPx Protection of 8-Aryl-dG Nucleosides

$N^2$-(Dimethylformamidyl)-8-aryl-2′-dG (1.67 mmol) was placed in a two-necked RBF, co-evaporated from dry pyridine (3 x 5 mL) and reverse-filled with argon. The RBF was then fitted with a dropping funnel, to which a DMPx-Cl (0.68 g, 1.91 mmol) pyridine (10 mL) solution was added. 10 mL of dry pyridine (or DMF for solubility considerations) was then added to the RBF, which was then placed in an ice bath and allowed to cool to 0°C. The DMPx-Cl/pyridine solution was then delivered to the RBF over 1 hour and the reaction was monitored by TLC. Upon completion, the mixture was diluted with ethyl acetate (10 mL) and washed with water (2 x 10 mL). TEA (1 mL) was added and the mixture was evaporated to dryness. The solid was then recrystallized by first dissolving in dichloromethane (3 mL), followed by the addition of hexanes (10 mL). The resulting suspension was filtered, and if the product was sufficiently pure it was carried forward; otherwise it was loaded onto a silica column and eluted with MeOH:CH$_2$Cl$_2$:TEA (5:90:5) to afford the product.

5′-O-DMPx-$N^2$-(Dimethylformamidyl)-8-(2′′-Furanyl)-2′-dG (5a).

Synthesis performed as outlined afforded 0.76 g (1.13 mmol) as an off white powder corresponding to a 68% yield. MP = 225–227°C. $^1$H NMR (600 MHz, CD$_2$Cl$_2$) $\delta = 8.99$ (bs, 1H), 8.12 (s, 1H), 7.44 (d, $J = 1.0$ Hz, 1H), 7.28-6.87 (m, 12H), 6.63 (dd, $J = 6.0$ Hz, 7.8 Hz, 1H), 6.50 (m, 1H), 4.55, (m, 1H), 4.14 (m, 1H) 3.45 (m, 1H), 3.31 (m, 1H), 3.15 (m, 1H), 3.05 (s, 3H), 2.86 (s, 3H), 2.26 (m, 1H), 2.19 (s, 3H), 1.84 (s, 3H); $^{13}$C NMR (600 MHz, CD$_2$Cl$_2$) $\delta = 157.84$, 157.81, 156.2, 151.1, 149.7, 149.6, 149.4, 145.3, 144.0, 140.1, 133.22, 133.17, 130.4, 130.3, 129.5, 129.3, 128.2, 126.9, 126.4, 122.9, 122.8, 121.2, 116.5, 116.1, 112.9, 112.1, 86.3, 84.7, 76.2, 73.3, 65.3, 46.5, 41.7, 39.0, 35.3, 20.9, 20.5. HRMS calcd for C$_{39}$H$_{37}$N$_6$O$_6$ $^{+}$ [M+H$^+$] 673.2776, found 673.2775.
5′-O-DMPx-N²-(Dimethylformamidyl)-8-Phenyl-2′-dG (5b).

Synthesis performed as outlined afforded 0.68 g (1.0 mmol) as a white powder corresponding to a 60% yield. MP = 195–198°C. ¹H NMR (600 MHz, CD₂Cl₂) δ = 9.11 (bs, 1H), 8.12 (s, 1H), 7.70 (d, J = 7 Hz, 2H), 7.43-7.38 (m, 3H), 7.31 (d, J = 7.4 Hz, 2H), 7.21 (t, J = 7.4 Hz, 2H), 7.13-7.08 (m, 2H), 7.02 (m, 2H), 6.93 (m, 3H), 6.24 (t, J = 6.4 Hz, 1H), 4.43 (m, 1H), 4.14 (m, 1H), 3.52 (t, J = 9.2 Hz, 1H), 3.34 (dd, J = 2.8, 9.4 Hz, 1H), 3.11 (m, 1H), 3.07 (s, 3H), 2.85 (s, 3H), 2.20 (s, 3H), 2.11 (s, 3H) 2.10 (m, 1H); ¹³C NMR (600 MHz, CD₂Cl₂) δ = 158.2, 158.0, 156.1, 149.9, 149.8, 149.7, 149.6, 133.4, 133.3, 130.9, 130.5, 130.4, 130.1, 129.7 129.5, 129.0, 128.4, 127.0, 126.5, 123.0, 121.0, 116.7, 116.2, 86.7, 85.0, 76.3, 73.4, 65.5, 41.9, 38.5, 35.5, 21.1, 20.7. HRMS calcd for C₄₀H₃₉N₆O₅+ [M+H⁺] 683.2982; found 683.2985.

5′-O-DMPx-N²-(Dimethylformamidyl)-8-(4″-Cyanophenyl)-2′-dG (5c)

Synthesis performed as outlined afforded 0.87 g (1.23 mmol) of a yellow powder corresponding to a 74% yield. MP = 170–173°C (d). ¹H NMR (600 MHz, DMSO-d₆) δ = 11.37 (bs, 1H), 7.99 (s, 1H), 7.95 (d, J = 8.2 Hz, 2H), 7.83 (d, J = 8.2 Hz, 2H), 7.19-7.04 (m, 7H), 6.87-6.83 (m, 2H), 6.75-6.72 (m, 2H), 6.05 (m, 1H), 5.24 (m, 1H), 4.40 (m, 1H), 4.03 (m, 1H), 3.4 (t, J = 9.5 Hz, 1H), 3.26 (m, 1H), 3.08 (m, 1H), 3.01 (s, 3H), 2.78 (s, 3H), 2.11 (s, 3H), 2.04 (m, 1H), 1.62 (s, 3H); ¹³C NMR (600 MHz, DMSO-d₆) δ = 157.6, 157.4, 156.4, 150.5, 149.3, 148.5, 148.4, 146.2, 134.6, 132.6, 132.2, 131.7, 130.1, 129.7, 129.5, 128.5, 128.4, 127.9, 126.4, 125.5, 122.3, 122.26, 120.7, 118.6, 116.2, 115.5, 111.7, 86.7, 84.6, 74.9, 71.3, 64.7, 40.9, 40.4, 37.1, 34.7, 20.4, 19.9. HRMS calcd for C₄₀H₃₉N₆O₅⁺ [M+H⁺] 708.2929; found 708.2932.

5′-O-DMPx-N²-(Dimethylformamidyl)-dG (5g).

Synthesis performed as outlined afforded 0.79 g (1.30 mmol) of a white powder corresponding to 78% yield. MP = 113–118°C (d). ¹H NMR (300 MHz, CDCl₃) δ = 9.50 (bs, 1H), 8.43 (s, 1H), 7.65 (s, 1H), 7.31 (m, 2H), 7.22-6.95 (m, 7H), 6.90 (m, 2H), 6.33 (t, J = 6.9 Hz, 1H), 4.55 (m, 1H), 4.13 (m, 1H), 3.16 (m, 2H), 3.02 (s, 3H), 2.99 (s, 3H), 2.54 (m, 2H), 2.13 (s, 3H), 2.08 (s, 3H); ¹³C NMR (300 MHz, CDCl₃) δ = 158.1, 158.0, 156.6, 150.1, 149.3, 149.28,
148.5, 132.8, 132.7, 130.2, 130.1, 129.3, 129.1, 127.9, 126.6, 126.3, 122.2, 122.0, 120.3, 116.0, 85.9, 83.3, 76.1, 72.5, 64.0, 45.7, 41.3, 40.8, 35.1, 30.9, 20.7, 20.6. HRMS calcd for C_{34}H_{35}N_{6}O_{5}^{+} [M+H^{+}] 607.2669; found 607.2672.

5′-O-DMPx-N^{4}-(Acetyl)-2′-dC (5h).

Synthesis performed as outlined afforded 0.86 g (1.55 mmol) of a white powder corresponding to a 93% yield. MP = 115–118°C (d). ¹H NMR (300 MHz, CDCl₃) δ = 9.03 (bs, 1H), 8.21 (d, J = 7.5 Hz, 1H), 7.36-7.11 (m, 10H), 6.92 (m, 2H), 6.22 (t, J = 6.1 Hz, 1H), 4.35 (m, 1H), 4.06 (m, 1H), 3.24 (dd, J = 3.2, 10.6 Hz, 1H), 3.10 (dd, J = 4.1, 10.6 Hz, 1H), 2.78-2.70 (m, 1H), 2.22 (m, 10H); ¹³C NMR (300 MHz, CDCl₃) δ = 170.6, 162.8, 155.6, 150.0, 149.9, 148.7, 144.6, 133.42, 133.37, 130.8, 129.5, 129.3, 128.3, 127.2, 126.8, 122.5, 122.2, 116.6, 96.3, 87.9, 87.0, 76.9, 71.9, 63.6, 42.4, 25.2, 21.0. HRMS calcd for C_{32}H_{32}N_{3}O_{6}^{+} [M+H^{+}]: 554.2294; found 554.2296.

5′-O-DMPx-2′-dT (5i).

Synthesis performed as outlined afforded 0.81 g (1.47 mmol) of a white powder corresponding to a 88% yield. MP = 139–141°C (d). ¹H NMR (600 MHz, DMSO-d₆) δ = 11.33 (s, 1H), 7.53 (s, 1H), 7.27-7.24 (m, 4H), 7.16 (m, 3H), 7.11 (m, 2H), 7.04 (bs, 1H), 6.97 (m, 1H), 6.20 (t, J = 7Hz, 1H), 5.32 (d, J = 4.4 Hz, 1H), 4.31 (m, 1H), 3.82 (m, 1H), 3.13 (dd, J = 2.8, 10.4 Hz, 1H), 3.06 (dd, J = 4.0, 10.4 Hz, 1H), 2.27 (m, 1H), 2.18 (m, 1H), 2.14 (s, 3H), 3.10 (s, 3H), 1.42 (s, 3H). ¹³C NMR (600 MHz, DMSO-d₆) δ = 163.6, 150.4, 149.0, 148.7, 148.4, 135.7, 132.5, 132.4, 130.4, 130.2, 128.3, 128.22, 128.18, 126.8, 125.6, 122.1, 121.9, 116.1, 109.5, 85.3, 83.6, 75.5, 70.8, 63.7, 20.3, 20.2, 11.7. HRMS calcd for C_{32}H_{32}N_{2}O_{6}Na⁺ [M+Na⁺]: 549.2002; Found 549.1995.

5′-O-DMPx-N^{6}-(Dimethylformamidyl)-dA (5j).

Synthesis performed as outlined afforded 0.83g product as a white powder in 84% yield. MP = 113-118°C. ¹H NMR (300 MHz CDCl₃) δ = 8.94 (s, 1H), 8.46 (s,1H), 8.02
(s, 1H), 7.34–7.31 (m, 2H), 7.27–7.17 (m, 3H), 7.07–7.04 (m, 4H), 6.94–6.92 (m, 2H), 6.44 (t, J = 6.6 Hz, 1H), 4.63 (m, 1H), 4.10 (m, 1H), 3.27–2.15 (m, 8H), 2.80 (m, 1H), 2.18 (s, 3H), 2.17 (s, 3H). $^{13}$C NMR (300 MHz CDCl$_3$) $\delta = 159.2, 158.1, 152.1, 150.9, 149.2, 149.1, 148.3, 139.8, 132.4, 132.3, 129.9, 129.1, 128.8, 127.6, 126.3, 126.2, 126.0, 122.0, 121.8, 115.8, 86.2, 84.4, 75.9, 71.7, 63.6, 41.0, 40.5, 34.8, 20.5. HRMS calcd for C$_{34}$H$_{35}$N$_6$O$_4^+$ [M+H$^+$]: 591.272; found 591.2724.

2.3.2.5. Phosphoramidite Reaction

Effective reaction of the 3'-hydroxyl group with the β-cyanoethylphosphoramidite was achieved using standard phosphoramidite chemistry, combined with an adapted method of purification. Formerly, after the reaction was completed, the reaction volume was reduced in vacuo and subsequently purified by column chromatography. Results were mixed, the main difficulty being the acquisition of a pure $^{31}$P NMR spectrum, as oxidation from phosphite to phosphate occurs readily upon exposure to air and the impurities are difficult to visualize during purification. This is characterized by a peak ~14 ppm in the $^{31}$P NMR spectrum for the oxidized phosphate, while the desired reactive phosphite provides peaks at ~147 and 148 for the diastereomeric mixture. Other impurities due to any remaining unreacted phosphitylating agent display a series of peaks at ~5 ppm and again are difficult to remove via column chromatography, as visualization and specific elution is difficult. A synthetic procedure from the literature was adapted and modified in order to improve the purity of analytical samples.$^{54}$ Following the reaction, the organic layer was washed with degassed 5% HCO$_3$\textsubscript{(aq)}, dried over Na$_2$SO$_4$ then recrystallized by the slow addition of the amidite in dry, degassed CH$_2$Cl$_2$ to hexanes at -78° C with vigorous stirring. This step seemed to remove any unreacted phosphitylating agent and was found to work most effectively when employed prior to the column chromatography. The amidite was then isolated by filtration under an inert atmosphere and, if needed, subjected to
column chromatography for final cleanup. This procedure was found to yield the amidites 6a-6j in good yields with purities suitable for DNA synthesis. Typically multiple columns were needed for isolation of analytically pure samples.

2.3.2.5.1. General Procedure

The 5′-O-DMT(DMPx)-N²-(dimethylformamidyl)-8-aryl-2′-dG (0.706 mmol) was co-evaporated from dry THF (3 x 5 mL), reverse-filled with argon and dissolved in 10 mL dry, degassed CH₂Cl₂. To this was added dry, degassed TEA (0.4 mL, 2.83 mmol) and 2-cyanoethyl N,N-diisopropylchlorophosphoramide (0.24 mL, 1.06 mmol). The reaction was monitored via TLC and, upon completion (20-40 min), was washed successively with saturated, degassed sodium bicarbonate solution. The organic phase was separated, dried with Na₂SO₄ and purified by a combination of crystallization, by initially dissolving in methylene chloride and dripping into hexane at -78°C, and flash chromatography, by loading onto a flash chromatography column eluting with 92:5:3 CH₂Cl₂:MeOH:TEA. Phosphoramidites were isolated as their corresponding diastereomers, which were typically off-white foams.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N²-(Dimethylformamidyl)-8-(2′″-Furanyl)-2′-dG (4a).

The phosphoramidite 5a eluted from column chromatography as its diastereomers, which were a white foam (241mg, 76.0%): ¹H NMR (600 MHz, CDCl₃) δ = 9.28–9.24 (m, 1H), 8.32–8.27 (m, 1H), 7.44–7.06 (m, 13H), 6.66–6.56 (m, 5H), 6.43 (bs, 1H), 5.03–4.90 (m, 1H), 4.12 (m, 1H), 3.73–3.62 (m, 7H), 3.53–3.42 (m, 3H), 3.34–3.20 (m, 3H), 2.99 (s, 3H), 2.90 (s, 3H), 2.69–2.50 (m, 2H), 2.40–2.27 (m, 2H), 1.46 (m, 2H), 1.11–1.05 (m, 9H), 1.00–0.98 (m, 3H); ¹³C NMR (151 MHz, CDCl₃) δ = 158.32, 158.3, 157.7, 157.6, 157.5, 155.8, 155.8, 150.52, 150.50, 144.82, 144.80, 144.7, 143.60, 143.55, 140.39, 140.35, 135.82, 135.76, 135.7, 130.0, 129.9, 129.87, 128.1, 128.0, 127.7, 127.6, 126.7, 126.6, 126.6, 121.1, 121.0, 117.6, 117.5, 112.93,
112.88, 112.85, 112.8, 111.73, 111.71, 86.04, 85.99, 84.9, 84.85, 84.78, 84.2, 83.9, 74.7, 74.5, 73.5, 73.4, 63.8, 63.3, 58.4, 58.3, 58.1, 58.0, 57.9, 55.2, 55.1, 46.7, 45.2, 43.3, 41.1, 37.5, 35.1, 30.9, 29.2, 24.6, 24.5, 24.4, 24.3, 22.9, 22.8, 20.4, 20.2, 20.1. $^{31}$P NMR (121.4 MHz, CDCl$_3$) δ 149.32, 148.93; HRMS calcd for C$_{47}$H$_{56}$N$_8$O$_8$P$^+$ [M + H$^+$] 891.3959, found 891.3964.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-DMPx-N$^2$-Dimethylformamide-8′-(2″-Furanyl)-2′-dG (6a).

The phosphoramidite 6a was isolated as its diastereomers which were an off-white foam in 72% yield. $^1$H NMR (300 MHz, CD$_2$Cl$_2$) δ = 8.21 (m, 1H), 7.54 (m, 1H), 7.42-6.87 (m, 12H), 6.67 (m, 1H), 6.56 (m, 1H), 4.81 (m, 1H), 4.27 (m, 1H), 3.80-3.34 (m, 6H), 3.19-3.10 (m, 4H), 2.92 (m, 3H), 2.56 (m, 2H), 2.21 (m, 3H), 1.89 (m, 3H), 1.37 (m, 2H), 1.19-1.09 (m, 12H). $^{13}$C NMR (300 MHz, CD$_2$Cl$_2$) δ = 158.0, 157.9, 156.3, 151.2, 149.8, 149.7, 148.6, 145.5, 144.1, 140.3, 133.4, 133.3, 130.5, 130.4, 129.6, 129.5, 127.0, 123.0 122.9, 121.3, 117.7, 116.2, 113.0, 112.2, 86.5, 84.8, 76.3, 73.4, 65.4, 57.9, 46.6, 43.3, 41.9, 39.1, 35.4, 23.4, 23.2, 21.0, 20.6, 20.2. $^{31}$P NMR (300 MHz, CDCl$_3$) δ = 149.3, 149.0. HRMS calcd for C$_{47}$H$_{54}$N$_8$O$_7$P$^+$ [M+H$^+$] 873.3848; found 873.3852.

3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-ODMT-N$^2$-(Ddimethylformamidyl)-8-(Phenyl)-2′-dG (4b).

Product isolated in 83% yield as corresponding diastereomers which were a foam. $^1$H NMR (300 MHz, CDCl$_3$) δ = 9.17 (bs, 1H), 8.35 (m, 1H), 7.80 (m, 2H), 7.37-7.07 (m, 13H), 6.65 (m, 4H), 6.14 (m, 1H), 5.06-4.90 (m, 1H), 4.08 (t, J = 4.8 Hz, 1H), 3.71-3.56 (m, 7H), 3.51-3.17 (m, 7H), 2.99 (s, 3H), 2.85 (s, 3H), 3.44 (m, 1H), 2.35 (m, 3H), 1.08 (m, 6H), 0.97 (m, 6H); $^{13}$C NMR (300 MHz, CDCl$_3$) δ = 157.9, 157.8, 157.54, 157.52, 156.3, 150.4, 148.2, 145, 135.66, 135.6, 130.3, 129.6, 129.44, 129.40, 129.2, 128.6, 127.60, 127.59, 126.5, 120.1, 112.93, 112.89,
85.8, 85.1, 84.1, 70.9, 64.0, 58.7, 54.9, 54.9, 43.7, 40.8, 40.0, 37.3, 34.6, 23.3, 23.2, 23.1, 20.4, 20.3; $^{31}$P NMR $\delta = 149.32, 148.93$

3′-O-[2-Cyanoethoxy](diisopropylamino)phosphino]-5′-O-DMPx-N$^2$-(Dimethylformamidyl)-8-(Phenyl)-2′-dG (6b).

Product isolated in 80% yield as corresponding diastereomers which were a foam. $^1$H NMR (300 MHz, CDCl$_3$) $\delta = 9.16$ (bs, 1H), 8.19-8.11 (m, 1H), 7.61-6.86 (m, 16H), 6.13 (m, 1H), 4.52 (m, 1H), 4.11 (m, 1H), 3.90 (m, 1H), 3.80-3.51 (m, 4H), 3.46 (s, 3H), 3.43 (s, 3H), 2.68-2.50 (m, 3H), 2.15 (m, 1H), 2.03 (s, 3H), 2.00 (s, 3H), 1.05-0.99 (m, 12H); $^{13}$C NMR (300 MHz, CDCl$_3$) $\delta = 158.4, 158.2, 156.3, 151.8, 150.1, 150.0, 149.9, 149.8, 133.5, 133.5, 131.1, 130.7, 130.6, 130.3, 129.8, 129.7, 129.2, 128.6, 127.2, 126.7, 123.2, 123.2, 121.2, 116.9, 116.4, 86.9, 85.1, 76.5, 73.6, 65.7, 58.5, 43.9, 43.8, 42.1, 38.7, 35.6, 23.1, 23.0, 22.9, 21.3, 20.9, 20.2; $^{31}$P NMR $\delta = 148.74, 148.67$. HRMS calcd for C$_{49}$H$_{56}$N$_8$O$_6$P $[M+H]^+$ 883.4055; found 883.4060.

3′-O-[2-Cyanoethoxy](diisopropylamino)phosphino]-5′-O-DMT-N$^2$-(Dimethylformamidyl)-8-(4′′-Cyanophenyl)-2′-dG (4c).

Product isolated in 68% yield as corresponding diastereomers which were a foam. $^1$H NMR (400 MHz, CDCl$_3$) $\delta = 9.05-8.98$ (m, 1H), 8.41-8.34 (m, 1H), 8.05-8.02 (m, 2H), 7.72-7.67 (m, 2H), 7.39-7.36 (m, 2H), 7.26-7.13 (m, 8H), 6.73-6.67 (m, 4H), 6.13-6.11 (m, 1H), 5.21-4.91 (m, 1H), 4.17-4.14 (m, 1H), 3.70-3.31 (m, 12H), 3.05 (s, 3H), 2.89 (s, 3H), 2.52-2.33 (m, 3H), 1.39-1.38 (m, 1H), 1.13-1.01 (m, 12H); $^{13}$C NMR (400 MHz, CDCl$_3$) $\delta = 158.43, 158.38, 157.6, 151.20, 151.15, 147.6, 144.7, 144.6, 135.83, 135.76, 135.7, 134.4, 132.21, 132.15, 130.13, 130.08, 130.02, 129.97, 129.91, 128.12, 128.06, 127.8, 127.7, 126.8, 126.7, 121.3, 118.4, 117.5, 113.0, 112.96, 112.9, 86.1, 84.2, 83.8, 63.4, 62.7, 58.4, 55.2, 55.1, 45.3, 45.2, 43.3, 43.25, 41.2, 37.1, 35.2, 31.5, 24.6, 24.5, 24.3, 22.9, 22.8, 22.6, 20.3, 20.1, 19.2, 14.1; $^{31}$P NMR (300 MHz, CDCl$_3$) $\delta = 149.38, 149.06$. HRMS calcd for C$_{50}$H$_{57}$N$_9$O$_7$P $[M+H]^+$ 926.4113; found 926.4117.
3'-O-[2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N²-(Dimethylformamidyl)-8-(4''-Cyanophenyl)-2'-dG (6c).

Product isolated in 65% yield as corresponding diastereomers which were a foam. ¹H NMR (300 MHz, CD₃OD) δ = 8.21 (s, 1H), 7.95-7.84 (m, 4H), 7.29-6.78 (m, 11H), 6.15 (m, 1H), 4.95 (m, 1H), 4.37-4.27 (m, 1H), 4.22-4.09 (m, 2H), 3.84-3.47 (m, 7H), 3.16 (m, 3H), 2.97 (m, 3H), 2.87 (t, J = 5.9 Hz, 2H), 2.16 (m, 3H), 1.75 (m, 3H), 1.30 (d, J = 5.2 Hz, 6H), 1.27 (d, J = 5.2 Hz, 6H); ¹³C NMR (300 MHz, CD₃OD) δ = 160.3, 159.0, 158.1, 152.5, 151.5, 150.8, 150.5, 149.0, 135.4, 133.9, 133.7, 133.4, 131.4, 131.2, 130.9, 130.6, 128.8, 127.6, 127.3, 124.0, 123.8, 131.7, 119.3, 117.3, 116.8, 114.5, 87.5, 86.4, 77.2, 70.0, 65.9, 60.3, 59.7, 59.5, 46.7, 46.6, 44.7, 44.6, 44.5, 42.0, 35.6, 25.1, 25.0, 23.2, 20.8, 20.3; ³¹P NMR δ = 149.40, 149.12. HRMS calcd for C₅₀H₅₅N₉O₆P⁺ [M+H⁺] 908.4007; found 908.4013.

3'-O-[2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-N²-(Dimethylformamidyl)-8-(8''-Quinolyl)-2'-dG (4d).

Product isolated in 74% yield as corresponding diastereomers which were a foam. ¹H NMR (300 MHz DMSO-d₆) δ = 11.40 (bs, 1H), 8.92 (m, 1H), 8.53 (m, 2H), 8.26-8.19 (m, 2H), 8.00 (bs, 1H), 7.77 (m, 1H), 7.66 (dd, J = 4.2Hz, 8.3 Hz, 1H), 7.29-6.73 (m, 13H), 5.71 (bs, 1H), 5.10 (d, J = 4.5Hz, 1H), 4.41 (bs, 1H), 3.73-3.68 (m, 7H), 3.64 (m, 1H), 3.23-3.18 (m, 3H), 3.06-2.94 (s, 7H), 2.12 (bs, 1H), 1.06 (m,1H), 0.97-0.81 (m, 12H); ¹³C (300Mhz, CDCl₃) δ = 154.5, 154.4, 154.0, 147.6, 146.7, 141.5, 133.2, 132.1, 132.0, 129.2, 127.0, 126.1, 125.9, 124.3, 124.1, 123.0, 118.6, 109.5, 109.4, 81.7, 60.9, 85.2, 51.5, 51.4, 43.3, 43.1, 42.1, 37.2, 36.5, 34.4, 31.1, 27.2, 23.7, 23.6, 20.9; ³¹P NMR δ = 149.32, 148.88. HRMS calcd for C₅₀H₅₅N₉O₆P⁺ [M+H⁺] 952.4275; found 952.4278.
3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N2-(Dimethylformamidyl)-8-(2′′-Benzothienyl)-2′-dG (4e).

The phosphoramidite 5a eluted from column chromatography as its diastereomers, which were a white foam (480 mg, 72.0%). \(^1\)H NMR (400MHz (CD$_3$)$_2$CO): 10.62 (bs, 1H), 8.60 (s, 1H), 8.07 (s, 1H), 7.94 (m, 1H), 7.87 (m, 1H), 7.41-7.07 (m, 11H), 6.74-6.65 (m, 5H), 5.39 (m, 1H), 4.30-4.23 (m, 1H) 3.68 (s, 3H), 3.66 (s, 3H), 3.63-3.57 (m, 2H), 3.43-3.31 (m, 3H), 3.23-3.18 (m, 3H) 3.18-2.91 (m, 5H), 2.90-2.78 (m, 2H), 2.66-3.58 (m, 1H). \(^13\)C NMR (400MHz (CD$_3$)$_2$CO) 159.4, 159.3, 159.2, 158.7, 157.7, 157.1, 151.9, 146.1, 143.9, 141.1, 140.9, 136.8, 136.7, 130.8, 130.6, 128.9, 128.8, 128.5, 128.4, 127.4, 126.3, 125.55, 125.5, 125.3, 124.8, 122.8, 119.2, 118.8, 113.73, 113.7, 86.6, 86.0, 85.9, 85.6, 85.5, 75.0, 74.7, 64.2, 64.0, 59.7, 59.6, 55.42, 55.35, 45.7, 45.7, 44.1, 44.0, 37.8, 37.6, 35.3, 35.2, 24.9, 24.8, 23.1, 23.0, 20.3, 20.2. \(^{31}\)P NMR (300MHz (CD$_3$)$_2$CO) \(\delta = 148.3, 148.0.\) HRMS calcd for C$_{51}$H$_{58}$N$_8$O$_7$PS+ [M+H$^+$] 957.3881; found 957.3879.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N2-(Dimethylformamidyl)-8-(Pyrenyl)-2′-dG (4f).

Reaction produced 0.60 g of product corresponding to an 84% yield of the diastereomers which were a foam. \(^1\)H NMR (600MHz, CD$_2$Cl$_2$) \(\delta = 9.06 (bs, 1H), 8.52 (m, 1H), 8.29-8.09 (m, 9H), 7.41 (m, 2H), 7.27-7.21 (m, 7H), 6.76 (m, 4H), 6.06 (bs, 1H), 3.93, (bs, 1H), 3.76 (s, 3H), 3.75 (s, 3H), 3.41 (bs, 1H), 3.36-3.25 (m, 4H), 3.20 (bs, 1H), 3.11 (bs, 1H), 3.08 (m, 3H), 3.0 (m, 3H), 2.28-2.21 (m, 3H), 1.02 (s, 3H), 0.98 (d, 3H), 0.84 (d,3H), 0.79 (d,3H). \(^{13}\)C NMR (600MHz, CD$_2$Cl$_2$): 158.6, 157.9, 157.6, 151.0, 145.1, 145.0, 136.0, 135.8, 132.4, 131.3, 130.9, 130.1, 130.0, 128.9, 128.7, 128.6, 128.1, 127.8, 124.6, 124.4, 124.35, 117.64, 113.1, 86.11, 84.5, 80.1, 77.0, 64.1, 63.6, 58.4, 57.9, 55.3, 43.2, 43.0, 41.1, 37.4, 35.0, 24.2, 24.1, 20.3, 20.1. \(^{31}\)P NMR
(300MHz, CD$_2$Cl$_2$): 149.2, 148.8. HRMS calcd for C$_{39}$H$_{62}$N$_8$O$_7$P$^+$ [M+H$^+$]: 1025.4474; found 1025.4490.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$^2$-
(Dimethylformamidyl)-dG (6g)

Product isolated in 91% yield as corresponding diastereomers which were a foam. $^1$H NMR (300 MHz, CDCl$_3$) δ = 9.70 (bs, 1H), 8.45 (m, 1H), 7.60 (m, 1H), 7.26-6.86 (m, 11H), 6.21 (m, 1H), 4.58 (m, 1H), 4.16 (m, 1H), 3.74-3.37 (m, 4H), 3.12 (m, 2H), 3.01 (s, 6H), 2.43-2.36 (m, 4H), 2.07 (m, 6H), 1.11-1.00 (m, 12H); $^{13}$C NMR (300 MHz, CDCl$_3$) δ = 161.1, 161.0, 159.6, 153.0, 152.3, 152.3, 151.5, 135.8, 135.7, 133.2, 133.1, 132.3, 132.1, 130.8, 129.6, 129.3, 125.1, 125.0, 123.2, 119.0, 117.7 88.9, 86.3, 79.1, 75.4, 67.0, 57.2, 44.3, 43.8, 42.8, 38.0, 33.9, 23.74 23.68, 22.7, 19.6. $^{31}$P NMR δ = 148.9, 148.7. HRMS calcd for C$_{43}$H$_{52}$N$_8$O$_6$P$^+$ [M+H$^+$]: 807.3748; found 807.3752.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-N$^4$-Ac-dC (6h).

Product isolated in 91% yield as corresponding diastereomers which were a foam. $^1$H NMR (300 MHz CDCl$_3$) δ = 1H 10.25 (bs, 1H), 8.22-8.05 (m, 1H), 7.33-6.78 (m, 12H), 6.17 (m, 1H), 4.32 (m, 1H), 4.09 (m, 1H), 3.73-3.41 (m, 4H), 3.17 (m, 1H), 2.98 (m, 1H), 2.70 (m, 1H), 2.52 (t, J = 6.3 Hz, 2H), 2.32 (m, 1H), 2.22 (s, 3H), 2.16 (m, 7H), 0.95 (m, 12H); $^{13}$C NMR (300 MHz CDCl$_3$) δ = 169.1, 168.9, 161.4, 154.1, 148.4, 147.1, 143.1, 131.9, 129.3, 128.0, 126.8, 125.6, 125.3, 121.0, 120.7, 117.1, 115.0, 113.8 94.8, 86.4, 85.4, 75.4, 70.3, 62.1, 58.1, 43.4, 43.4, 40.8, 23.7, 23.6, 23.2, 23.1, 20.6, 20.5, 19.5, 19.4; $^{31}$P NMR δ = 149.6, 149.2. HRMS calcd for C$_{41}$H$_{69}$N$_8$O$_7$P$^+$ [M+H$^+$]: 754.3369 found 754.3373.
3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-T (6i)

Product isolated in 93% yield as corresponding diastereomers which were a foam. $^1$H NMR (400 MHz CD$_2$Cl$_2$) $\delta =$ 9.22 (bs, 1H), 7.67 (m, 1H), 7.357.06 (m, 10H), 6.93 (m, 1H), 6.35 (m, 1H), 4.55 (m, 1H), 4.13-4.07 (m, 1H), 3.63-3.55 (m, 4H), 3.30-3.27 (m, 1H), 3.17-3.14 (m, 1H), 2.53 (m, 1H), 2.41-2.33 (m, 3H), 2.24 (s, 3H), 2.24 (s, 3H), 2.18 (s, 3H), 1.63 (s, 3H), 1.23 (12H), 1.18 (d, J = 6.8 Hz, 12H); $^{13}$C NMR (400 MHz CD$_2$Cl$_2$) $\delta =$ 164.8, 151.1, 150.0, 149.6, 149.0, 136.1, 133.3, 130.8, 130.7, 129.3, 129.0, 128.4, 127.2, 126.6, 122.4, 122.1, 118.2, 116.5, 111.4, 85.6, 85.0, 76.9, 74.8, 74.6, 63.7, 60.7, 58.8, 58.6, 43.6, 40.4, 24.7, 20.8, 14.3, 12.3, 1.12; $^{31}$P NMR $\delta =$ 149.2, 148.3. HRMS calcd for C$_{40}$H$_{48}$N$_4$O$_7$P$^+ [M+H$^+$]: 727.3260; found 727.3264.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMPx-Ac-dC Amidite (6j)

Product isolated in 91% yield as corresponding diastereomers which were a foam. $^1$H NMR (300 MHz CDCl$_3$) 8.94 (s, 1H), 8.39-8.37 (m,1H), 8.03-8.00 (m,1H), 7.34-6.95 (m,11H), 6.40 (t, J = 6.5 Hz, 1H), 4.76 (m, 1H), 4.24 (m, 1H), 3.83-3.62 (m, 4H), 3.33-3.27 (m, 1H), 3.20-3.15 (m, 7H), 2.93 (m, 1H), 2.62 (m, 2H), 2.14 (m, 6H), 1.27-1.13 (m, 12H); $^{13}$C NMR (300 MHz CDCl$_3$) $\delta =$158.4, 157.3, 151.3, 150.0, 148.3, 148.25, 147.4, 138.9, 131.6, 131.5, 129.1, 128.2, 128.0, 126.7, 125.5, 125.3, 125.2, 121.1, 121.0, 114.9, 113.8, 85.7, 85.3, 75.1, 70.9, 65.7, 58.2, 58.17, 43.3, 40.1, 39.6, 34.0, 23.0, 20.4, 19.6; $^{31}$P NMR $\delta =$ 149.1, 148.8. HRMS calcd for C$_{43}$H$_{52}$N$_8$O$_7$P$^+ [M+H$^+$]: 791.3793 found 791.3773.

2.3.3. Synthesis of C8-Aryl-dG Modified Oligonucleotides

Synthesis of the C8—aryl-dG modified oligonucleotides were performed on a 1μmol scale on a BioAutomation Corporation MerMade12 automatic synthesizer using standard and modified $\beta$-cyanoethylphosphoramidite chemistry according to published protocols.$^{36}$ Details concerning the synthesis and purification of modified oligonucleotides are outlined in appendix A.
Following the removal of the last trityl group and oligonucleotide deprotection in ammonium hydroxide, the oligonucleotide solution was concentrated under diminished pressure using a ThermoSavant DNA 120 SpeedVac. Crude material was then re-dissolved in 18.2 MΩ H₂O, which was filtered using Mandel syringe filters (PVDF 0.20 μm) and subjected to HPLC purification. Purification was performed using an Agilent 1200 series HPLC instrument equipped with an autosampler, diode array detector (monitored at λₘₚₜ=258 nm and λₘₜₜ of the incorporated adduct), fluorescence detector (monitored at λₑₓ and λₑₘ of the incorporated adduct), and autocollector. For analytical sample sizes (injection volumes less than 100 μL), separations were carried out at 50°C using a Phenomenex Clarity 3 μm Oligo-RP C18 column (50 x 4.60 mm) with a flow rate of 0.5 mL/min. Semi-preparative samples (200 μL injection volumes) were separated at 50°C using a Phenomenox 5 μm Oligo-RP C18 column (100 x 10 mm) with a flow rate of 3.5 mL/min. In either case the elution gradient transitioned from “buffer A” to “buffer B” (Buffer A = 95:5 aqueous 50 mM TEAA (pH 7.2)/acetonitrile; buffer B=30:70 aqueous 50mM TEAA (pH 7.2)/acetonitrile). Lyophilisation of the isolated product was performed using a Labconco FreeZone 4.5. HPLC analysis of the final isolated product was used to confirm purity. Oligonucleotides were then dissolved into 18.2 MΩ water for quantification by UV-vis measurement. Extinction coefficients (ε₂₆₀) were obtained from the Integrated DNA Technologies (IDT) internet website (URL in Appendix A) using an established oligo analyzer program as previously reported; all modified oligonucleotides were assumed to have the same extinction coefficient at 260 nm as the unmodified sequence.⁵⁵

2.3.3.1. Mass Spectrometry Analysis

Samples were submitted for analysis by mass spectrometry to Aaron Witham at the University of Guelph, Guelph ON. MS experiments for DNA identification were conducted on a quadrupole ion trap SL spectrometer. Masses were acquired in the negative ionization mode with an electrospray ionization source. Oligonucleotide samples were prepared in 90% Milli-Q filtered
water/10% methanol containing 0.1 mM ammonium acetate. Full scan MS spectra were obtained by direct infusion at a rate of 5–10 μL/min. Peak assignments are shown in Table 2 - 1 and Table 2 - 2 with mass spectra presented in Appendix C.

Table 2 - 1: ESI-MS Analysis of 8-aryl-dG Modified NarI(12) oligonucleotides.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Product Formula</th>
<th>Calcd mass</th>
<th>Exptl m/z</th>
<th>Exptl Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>NarI(12) (X=Fur dG)</td>
<td>C₁₁₈H₁₄₉N₄₂O₇₂P₁₁</td>
<td>3646.6</td>
<td>[M-4H]⁺ = 910.7</td>
<td>3647.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-5H]⁻ = 728.4</td>
<td>3646</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-6H]⁻ = 607</td>
<td>3647</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-7H]⁻ = 520.2</td>
<td>3647.4</td>
</tr>
<tr>
<td>NarI(12) (X=Bth dG)</td>
<td>C₁₂₂H₁₅₂N₄₂O₇₄P₁₂S</td>
<td>3712.6</td>
<td>[M-6H]⁻ = 618</td>
<td>3712.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-7H]⁻ = 529.5</td>
<td>3712.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]⁻ = 463.2</td>
<td>3712.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]⁻ = 411.6</td>
<td>3713</td>
</tr>
<tr>
<td>NarI(12) (X=CNPh dG)</td>
<td>C₁₂₁H₁₅₀N₄₃O₇₁P₁₁</td>
<td>3681.7</td>
<td>[M-6H]⁻ = 612.8</td>
<td>3681.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-7H]⁻ = 525</td>
<td>3681</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]⁻ = 459.4</td>
<td>3682.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]⁻ = 408.2</td>
<td>3681.8</td>
</tr>
<tr>
<td>NarI (12) (X=Q dG)</td>
<td>C₁₂₃H₁₅₂N₄₃O₇₁P₁₁</td>
<td>3707.7</td>
<td>[M-6H]⁻ = 617.1</td>
<td>3707.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-7H]⁻ = 528.8</td>
<td>3707.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]⁻ = 462.6</td>
<td>3707.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]⁻ = 411.1</td>
<td>3707.9</td>
</tr>
<tr>
<td>NarI(12) (X=Py dG)</td>
<td>C₁₃₀H₁₅₇N₄₂O₇₁P₁₁</td>
<td>3782.7</td>
<td>[M-4H]⁺ = 944.8</td>
<td>3782.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-5H]⁻ = 755.7</td>
<td>3782.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-6H]⁻ = 629.6</td>
<td>3782.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-7H]⁻ = 539.5</td>
<td>3782.5</td>
</tr>
</tbody>
</table>
Table 2 - 2: ESI-MS Analysis of 8-aryl-dG Modified NarI(22) oligonucleotides

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Product Formula</th>
<th>Calcd mass</th>
<th>Exptl m/z</th>
<th>Exptl Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>NarI(22) (X=Fur) dG</td>
<td>C_{214}H_{271}N_{78}O_{131}P_{21}</td>
<td>6695.14</td>
<td>[M − 9H]^9 = 743.0</td>
<td>6696</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 10H]^10 = 668.6</td>
<td>6696</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 11H]^11 = 607.7</td>
<td>6695.7</td>
</tr>
<tr>
<td>NarI(22) (X=Ph) dG</td>
<td>C_{216}H_{273}N_{78}O_{131}P_{21}</td>
<td>6705.16</td>
<td>[M − 9H]^9 = 744.3</td>
<td>6705.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 10H]^10 = 669.6</td>
<td>6706</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 11H]^11 = 608.6</td>
<td>6705.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 12H]^12 = 557.8</td>
<td>6705.6</td>
</tr>
<tr>
<td>NarI(22) (X=Q) dG</td>
<td>C_{219}H_{274}N_{79}O_{131}P_{21}</td>
<td>6756.2</td>
<td>[M − 9H]^9 = 749.7</td>
<td>6756.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 10H]^10 = 674.7</td>
<td>6757</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 11H]^11 = 613.2</td>
<td>6756.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 12H]^12 = 562.0</td>
<td>6756</td>
</tr>
<tr>
<td>NarI(22) (X=Py) dG</td>
<td>C_{226}H_{279}N_{78}O_{131}P_{21}</td>
<td>6831.2</td>
<td>[M − 9H]^9 = 758.3</td>
<td>6831.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 10H]^10 = 682.3</td>
<td>6831</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 11H]^11 = 620.2</td>
<td>6831.2</td>
</tr>
</tbody>
</table>

2.4. Results and Discussion:

To test the utility of the 5′-O-DMPx group and draw comparison to 5′-O-DMT protection, four representative 8-aryl-dG adducts were synthesized using Suzuki cross-coupling procedures, and were then converted into phosphoramidites using standard protocols (Scheme 2 - 5). The nucleoside adducts chosen for study serve different purposes. The 5-membered heteroaryl-dG analog (Fur-dG) has been employed to study probe conformation in duplex DNA.\textsuperscript{7} Furan-decorated nucleobase analogs serve as efficient fluorescent nucleobase mimics\textsuperscript{56,57} and are precursors for oxidative-promoted DNA cross-link formation.\textsuperscript{58} The phenyl derivative (Ph-dG) is an authentic DNA adduct produced by arylhydrazine carcinogens.\textsuperscript{2,3} Attachment of the electron-withdrawing cyano group to afford CNPh-dG generates a nucleoside with donor-acceptor character. Biphenyl systems with D-A character produce charge-transfer (CT) states in the emission spectra.
that are very sensitive to solvent polarity,\textsuperscript{59} thereby permitting the use of fluorescence spectroscopy to characterize adduct conformation within duplex DNA. The bulky 8-quinoyl-dG derivative Q\textsubscript{dG} is a model for C-linked naphthalene-dG adducts that, when compared to Ph\textsubscript{dG}, permits assessment of aryl ring size on mutagenicity. The electron-withdrawing ring nitrogen in Q\textsubscript{dG} also provides D-A character.

The nucleoside adducts (Fur\textsubscript{dG}, Ph\textsubscript{dG}, CNPh\textsubscript{dG} and Q\textsubscript{dG}) were shown to possess very different sensitivity to acid-catalyzed deglycosylation (Table 2 - 3). First order rate constants (k\textsubscript{obs}) were determined using UV-vis by monitoring in 0.1 M HCl at 37\(^\circ\)C the appearance of a peak corresponding to the deglycosylated nucleobase as a function of time, as previously described.\textsuperscript{34,60} Under these conditions, dG undergoes hydrolysis with a half-life (t\textsubscript{1/2}) of \~ 18 min.\textsuperscript{22} Surprisingly, the Fur\textsubscript{dG} adduct was the most sensitive to acid and was hydrolyzed 55 times faster than dG. Interestingly, the bulky quinoline adduct (Q\textsubscript{dG}) was the least sensitive to acid with a hydrolysis rate only 5.6-fold faster than dG. This suggested that preferential protonation of the quinoline ring N atom (pK\textsubscript{a} 4.9\textsuperscript{21}), as opposed to N\textsubscript{7} of the nucleobase, helps protect the nucleoside from acid-catalyzed hydrolysis. Since Q\textsubscript{dG} was not particularly sensitive to acid, its phosphoramidite containing 5\textsuperscript{′}-O-DMPx was not prepared.

Table 2 - 3: Half-lives (t\textsubscript{1/2}) for hydrolysis of 8-aryl-dG adducts

<table>
<thead>
<tr>
<th>adduct</th>
<th>k\textsubscript{obs} (min\textsuperscript{-1}), t\textsubscript{1/2} (min)\textsuperscript{a}</th>
<th>k\textsubscript{obs}/k\textsubscript{obs(dG)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fur\textsubscript{dG}</td>
<td>2.16 ± 0.08, 0.321</td>
<td>55.2</td>
</tr>
<tr>
<td>Ph\textsubscript{dG}</td>
<td>0.79 ± 0.008, 0.877</td>
<td>20.2</td>
</tr>
<tr>
<td>CNPh\textsubscript{dG}</td>
<td>1.78 ± 0.02, 0.389</td>
<td>45.5</td>
</tr>
<tr>
<td>Q\textsubscript{dG}</td>
<td>0.22 ± 0.002, 3.16</td>
<td>5.6</td>
</tr>
<tr>
<td>dG</td>
<td>0.0391, 17.7\textsuperscript{b}</td>
<td>1</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Determined in 0.1 M HCl at 37\(^\circ\)C from the average of six kinetic runs. \textsuperscript{b} Data for dG taken from ref 8.
**Optimization of DNA Synthesis.** For optimization of solid-phase DNA synthesis, the most acid-sensitive $\text{Fur}dG$ was utilized. Initial experiments examined hydrolysis rates of $\text{Fur}dG$ (1a), its $N^2$-dimethylformamidyl derivative 2a, and the $5'$-O-DMT analog (3a) (Scheme 2 - 5) in DCA as a function of acid concentration (3%, 1% and 0.5% DCA) and solvent (H$_2$O vs. CH$_3$CN). The rate data (Table 2 - 4) demonstrated the impact of DCA concentration on hydrolysis rate and the need for dry solvents, as rates in H$_2$O were ~ 5-fold faster than rates in CH$_3$CN. For the $5'$-O-DMT analog (3a) a $t_{1/2}$ value of 17 min was determined in CH$_3$CN with 0.5% DCA at 21ºC, while the corresponding value in 3% DCA was 0.9 min (19.5-fold reduction in rate of hydrolysis). The rate data also demonstrated the protecting effect of the $N^2$-dimethylformamidyl group in CH$_3$CN. In both 1% and 0.5% DCA, hydrolysis rates for 2a were slower compared to the unprotected $\text{Fur}dG$ (1a).

Table 2 - 4: Summary of first-order rate constants ($k_{obs}$) and half-lives ($t_{1/2}$) for hydrolysis of $\text{Fur}dG$ and its derivatives

<table>
<thead>
<tr>
<th>adduct</th>
<th>solvent</th>
<th>3% DCA$^b$</th>
<th>1% DCA$^b$</th>
<th>0.5% DCA$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>H$_2$O$^a$</td>
<td>0.40, 1.7</td>
<td>0.21, 3.3</td>
<td>0.12, 5.9</td>
</tr>
<tr>
<td></td>
<td>CH$_3$CN$^a$</td>
<td>0.35, 2.0</td>
<td>0.15, 4.6</td>
<td>0.06, 11.4</td>
</tr>
<tr>
<td>2a</td>
<td>H$_2$O$^a$</td>
<td>0.84, 0.82</td>
<td>0.40, 1.7</td>
<td>0.22, 3.15</td>
</tr>
<tr>
<td></td>
<td>CH$_3$CN$^a$</td>
<td>0.85, 0.81</td>
<td>0.114, 6.1</td>
<td>0.042, 16</td>
</tr>
<tr>
<td>3a</td>
<td>H$_2$O$^a$</td>
<td>0.91, 0.76</td>
<td>0.40, 1.7</td>
<td>0.22, 3.15</td>
</tr>
<tr>
<td></td>
<td>CH$_3$CN</td>
<td>0.80, 0.9</td>
<td>0.097, 7.1</td>
<td>0.041, 17</td>
</tr>
</tbody>
</table>

$^a$ Contains 0.25% DMSO$^b$ $k_{obs}$ (min$^{-1}$), $t_{1/2}$ (min) determined by UV-vis at 21ºC from the average of six kinetic runs; errors < 5%.

It was determined that 3% DCA was indeed required to efficiently detritylate $5'$-O-DMT. Figure 2 - 10 depicts the differences between the rates of detritylation for $N^2$-dimethylformamidyl-$5'$-O-DMT-8-(2"-furanyl)-2'-dG in (i) 3% DCA and (ii) 0.5% DCA in acetonitrile. Upon detritylation the DMT carbocation appears immediately in 3%, but not in 0.5% DCA. These data suggested that the utility of the $5'$-O-DMPx protecting group, together
with a deblocking solution of 0.5% DCA in dry DCM, could enable efficient incorporation of Fur-dG into oligonucleotides, as hydrolysis was clearly limited in 0.5% DCA (Figure 2 - 10 (iii)).

![Graph showing differences between the rates of detritylation for 5'-O-DMT-N^2-dimethylformamidyl-8-(2''-furanyl)-2'-dG in 3% DCA and 0.5% DCA in acetonitrile.]

Figure 2 - 10: Differences between the rates of detritylation for 5'-O-DMT-N^2-dimethylformamidyl-8-(2''-furanyl)-2'-dG in (i) 3% DCA and (ii) 0.5% DCA in acetonitrile.

To test the efficiency of solid-phase DNA synthesis, phosphoramidites 6a and 4a (Scheme 2 - 5) were used to initially incorporate Fur-dG into a thymidine decamer (5'-T;XTT, X = Fur-dG). This model decamer was employed because thymidine is the most difficult nucleoside to detritylate\(^{22}\) and the easiest to convert into a phosphoramidite, since amine protection is not required. The reverse-phase (RP) HPLC profiles of the synthesis (1 μmol-scale) are shown in Figure 2 - 11.
Figure 2 - 11: HPLC traces for solid phase DNA synthesis of thymidine decamer (5'-T7Fur'dGTT) utilizing (top) DMT protected nucleosides; and (bottom) DMPx protected nucleosides. Red trace corresponds to fluorescence detection ($\lambda_{ex}$ 315 nm $\lambda_{em}$ 380 nm) while black trace corresponds to UV-Vis detection ($\lambda_{abs}$ = 260 nm).

For synthesis using 4a (Figure 2 - 11 (top)) and a deblock solution of 3% DCA in DCM (50 s detritylation time, continuous delivery to the column) the RP-HPLC trace suggested that full length decamer was prepared and eluted at ~ 22 min. This peak contained the characteristic fluorescent properties of the modified nucleobase Fur'dG (the dashed red trace is with fluorescence detection (FLD)). However, a dominant peak was also present at 13.6 min that lacked fluorescence. This peak resulted either from incomplete coupling of 4a or from acid-promoted depurination of Fur'dG, followed by hydrolysis of the abasic site during ammonolysis. To
determine if coupling efficiency was an issue, the truncated strand 5′-DMT-TXTT was synthesized. Detritylation was omitted in order to avoid exposure to acid. The HPLC trace can be seen below in Figure 2 - 12 and showed that coupling efficiency was not the problem, since the product contains both the Fur dG base, as evident from its fluorescence, and the subsequent dT, as evident from the long retention time that is due to the DMT group retained on the latter.

![HPLC Trace](image)

Figure 2 - 12: HPLC trace for solid phase DNA synthesis of (5′-DMT-T\textsuperscript{Fur}dGTT) utilizing standard protocols. Red trace corresponds to fluorescence detection (\(\lambda_{\text{ex}}\) 315 nm \(\lambda_{\text{em}}\) 380 nm) while black trace corresponds to UV-Vis detection (\(\lambda_{\text{abs}} = 260\) nm).

One interesting finding was that the Fur dG nucleobase was protected in an oligonucleotide strand that otherwise contained only cytidines. The HPLC trace shown below in Figure 2 - 13 had no deglycosylation artefacts and full length oligomer was formed. This finding can be explained by the pKa of protonated cytidine, which is 4.4 compared to the pKa of N\textsuperscript{7} protonated dG, which equals 2.34. Cytidine is preferentially protonated and acts like a “proton sponge” protecting the Fur dG from deglycosylation.
Figure 2 - 13: HPLC trace for the solid phase DNA synthesis of cytidine decamer (5′-CC7\textsuperscript{Fur}dGCC) using standard protocols.

**Synthesis of NarI Substrates**

To incorporate \textsuperscript{Fur}dG into the 12-mer NarI recognition sequence 5′-CTCGGCXCCATC (NarI(12)) using the 5′-O-DMPx phosphoramidites \textit{6a}, 5′-O-DMPx phosphoramidites of dG, dT and dC were utilized. Figure 2 - 14 displays the HPLC traces for the synthesis of NarI(12) using \textit{4a} and commercially available 5′-O-DMT phosphoramidites versus the synthesis using \textit{6a} and 0.5% DCA in DCM after the coupling of \textit{6a}.

The 1 µM synthesis of NarI(12) using \textit{4a} produced full length product, but contained peaks characteristic of failed hydrolysis (Figure 2 - 14 (top)), and after purification afforded a yield of 49 nmol for NarI(12). Using \textit{6a} and 5′-O-DMPx phosphoramidites of dG and dC, the hydrolysis peaks were eliminated (Figure 2 - 14 (b)) and an isolated yield of 190 nmol was obtained, for a 4-fold increase in yield of NarI(12). The 5′-O-DMPx phosphoramidites \textit{6b} and \textit{6c} (Scheme 2 - 5) were then employed to incorporate \textit{Ph}dG and \textit{CNPh}dG into the G\textsubscript{3}-site of NarI(12). A longer 22-mer substrate 5′-CTCGGCXCCATCCCTTACGAGC (NarI(22)) suitable for primer-extension assays using DNA polymerases was also prepared. For these NarI sequences ESI-MS spectra are available in Appendix C, ESI-MS analysis in Table 2 - 1 and isolated yields of oligonucleotide product are summarized in Table 2 - 5.
Table 2 - 5: Summary of yields for 8-aryl-dG modified NarI oligonucleotides

<table>
<thead>
<tr>
<th>ODN</th>
<th>X</th>
<th>Yield⁹⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td>NarI(12)</td>
<td>Fur dG⁹⁹</td>
<td>49 nmol</td>
</tr>
<tr>
<td></td>
<td>Fur dG</td>
<td>190 nmol</td>
</tr>
<tr>
<td></td>
<td>Ph dG</td>
<td>216 nmol</td>
</tr>
<tr>
<td></td>
<td>CNPh dG</td>
<td>170 nmol</td>
</tr>
<tr>
<td>NarI(22)</td>
<td>Fur dG</td>
<td>42 nmol</td>
</tr>
<tr>
<td></td>
<td>Ph dG</td>
<td>53 nmol</td>
</tr>
<tr>
<td></td>
<td>O dG⁹</td>
<td>26 nmol</td>
</tr>
</tbody>
</table>

⁹⁹ Isolated yield from 1 μmol-scale synthesis after oligonucleotides were purified via RP-HPLC

⁹ Abbrained nucleoside synthesized with 5′-O-DMT protection, while all other entries in table were synthesized using 5′-O-DMPx protection.

Figure 2 - 14: HPLC traces for solid phase DNA synthesis of NarI(12) (5′-CTCGGC⁹⁹FurGCCATC) utilizing (top) DMT protected nucleosides; and (bottom) DMPx protected nucleosides. Red trace corresponds to fluorescence detection (λex 315 nm λem 380 nm) while black trace corresponds to UV-Vis detection (λabs = 260 nm).
The results of our studies demonstrate the utility of the 5′-O-DMPx protecting group for solid-phase DNA synthesis of oligonucleotide substrates containing acid-sensitive 8-aryl-dG adducts. For the NarI substrates synthesized it was critical to remove the final 5′-OH protecting group on-column. Any attempts to utilize solid-phase extraction cartridges, with the final 5′-O-DMPx group attached to the oligonucleotide, resulted in degradation of the oligonucleotide through exposure of the 8-aryl-dG adduct to both acid and water. A number of examples in the literature have reported the use of solid-phase DNA synthesis for incorporation of 8-aryl-dG adducts into oligonucleotides using standard 5′-O-DMT phosphoramidites. For probe development, these include insertion of 8-pyridyl-dG,\(^{61}\) where preferential protonation of the pyridyl group (p\(K_a\) 5.2\(^{62}\)) would be expected to protect the base from depurination. In other examples the aryl ring is separated from the dG nucleobase by a vinyl\(^{61,63}\) or alkynyl group.\(^{64}\) This reduces steric strain between the aryl ring and sugar moiety, thereby diminishing the sensitivity of the nucleoside to acid-promoted hydrolysis.\(^7\) The 5′-O-DMPx protection strategy\(^{12}\) provides a more general method for insertion of 8-aryl-dG adducts into oligonucleotide substrates and does not significantly modify the protocols used for the traditional 5′-O-DMT solid-phase synthesis. In terms of DNA adduct formation by chemical mutagens, nucleophilic aryl radicals also attach to the 8-position of 2′-deoxyadenosine (dA).\(^{2,3,65}\) In terms of sensitivity to acid-catalyzed deglycosylation, dA is 2.2 times more reactive than dG at 30°C.\(^{23}\) No information is available on the hydrolytic stability of 8-aryl-dA adducts and very little has been reported on incorporation of 8-aryl-dA adducts into oligonucleotide substrates. A notable exception for probe development is the click-chemistry-based 8-triazole-dA,\(^{66}\) which again contains ring nitrogen atoms in the 8-aryl-substituent.

Chemical mutagens such as PAHs lack ring nitrogen atoms and attach directly to the 8-position of purine nucleosides, making them acid-sensitive. PAHs are ubiquitous environmental pollutants and are present in cigarette smoke and in vehicle exhaust condensate.\(^{67}\) Currently, very little is known about the biological implications of C-linked 8-purine adducts produced by PAHs.
Utilizing the solid-phase assisted synthesis strategies reported herein we have prepared a variety of adducted NarI(22) substrates; these have been provided to Anne Verway, University of Guelph, Guelph ON in order to assess the differing propensities of 8-aryl-dG adducts to induce mutagenesis when the adducted DNA is replicated.

2.5. Conclusions

The results of our studies demonstrate the utility of 5′-O-2,7-dimethylpixyl (DMPx) as a protecting group for any acid-sensitive, biologically relevant nucleosides for which incorporation into oligonucleotides using solid phase synthesis is required. This is demonstrated through the synthesis of 8-aryl-dG phosphoramidites for solid-phase assisted synthesis of oligonucleotide substrates. 8-Aryl-dG adducts are produced by a number of chemical mutagens that undergo bioactivation to afford aryl radical species that attach to the 8-position of purine nucleosides. These adducts also possess impressive fluorescent properties that make them desirable for use as fluorescent nucleobase probes. Unfortunately, 8-aryl-dG adducts are sensitive to acid-catalyzed deglycosylation, which limits the utility of standard solid-phase synthesis using 5′-O-DMT protection and a deblock solution consisting of 3% DCA in dry DCM. The 5′-O-DMPx group can be efficiently removed using 0.5% DCA in DCM, which limits exposure of the modified 8-aryl-dG lesion to acid during solid-phase synthesis. Using 8-furyl-dG (Fur-dG) as a representative acid-sensitive 8-aryl-dG adduct (55-fold more reactive than dG in 0.1 M HCl at 37°C), our results demonstrate that Fur-dG is much more stable in 0.5% DCA in CH₃CN versus 3% DCA (19.5-fold reduction in hydrolysis rate). This increased stability provides a 4-fold increase of yields in oligonucleotide solid-phase synthesis using 5′-O-DMPx phosphoramidites. These results demonstrate a real application for 5′-O-DMPx protection.
2.6. References


Chapter 3: Impact of 8-Aryl-Guanine Probes on DNA Duplex Structure
3.1. **BIG PICTURE**

Bulky DNA adducts are formed through the covalent attachment of aryl groups to the DNA nucleobases. Many of these adducts possess conformational heterogeneity, which has been proposed to give rise to different mutagenic outcomes. Some of the most important factors influencing the observed mutagenicity within bulky DNA adducts are the \textit{anti/syn} conformational preference of the adducts, their potential to inflict DNA mismatch stabilization, and their interactions with DNA polymerases and repair enzymes.

This chapter will focus on the structural implications that a family of C8-aryl-dG analogues have on duplex structure and various mismatch scenarios. The C8-aryl-dG analogues are structural mimics, not authentic standards of toxicologically relevant DNA adducts, and their fluorescent properties permit their use as diagnostic probes. The model compounds do resemble the authentic adducts with respect to the connectivity of the aromatic carbon skeleton attached at the C8 site. Modification at this site significantly impacts the \textit{syn/anti} conformational preferences of the nucleoside, a scenario shared between authentic DNA adducts and C8-aryl-dG adducts. These studies aim to dissect the correlation of adduct structure and mutagenic function, using a series of appropriately designed model compounds.

Relatively little information is available regarding C-linked C8-dG adducts compared to their N-linked counterparts, for which numerous authentic adducts have been synthesized to determine adduct impact within DNA. These studies should help to provide a systematic analysis to build on in coming years, where authentic C-linked C8-aryl-dG adducts will be incorporated into DNA structures to determine their mutagenic outcome.
3.2. Background

3.2.1. DNA Damage

A major aspect of DNA damage is the attachment of aryl and alkyl groups to DNA nucleobases to form DNA adducts.\(^1\) Some DNA adducts are small (e.g. abasic sites and oxidative adducts), while others are quite bulky (e.g. pyrimidine dimers, photoproducts, large carcinogen-bound adducts and cross-links).\(^2\) Many environmental contaminants have been identified as possible genotoxic agents that form bulky DNA adducts.\(^3\) Determination of exact mechanisms for the toxicity of these chemicals has been an active area of research. If the toxicity of such compounds is shown to arise from a mutagenic mechanism, more stringent legislations would be enforced, to limit the extent of exposure of humans via dietary intake, water and personal products.\(^4\)

An established principle of chemical carcinogenesis is that the covalent binding of carcinogens to DNA is causally related to tumorigenesis. This principle is supported by the following observations: (1) a majority of carcinogens are also mutagens; (2) the mutagenic and carcinogenic properties of many compounds depend on initial conversion or metabolism to electrophilic derivatives that react with nucleophilic sites of DNA; (3) mutagenic and carcinogenic responses are correlated with the extent of DNA adduct formation; and (4) the activation of certain proto-oncogenes can be accomplished through the interaction of carcinogens with DNA.\(^5\)

Typically, DNA adducts are formed in very low concentrations \textit{in vivo}, which makes their detection difficult. Methods developed in the mid 1980’s including various immunoassays\(^6\) and \(^{32}\)P-post-labelling techniques\(^7\) first provided the means to investigate adduct formation. Several theories have been proposed to understand how bulky adducts cause DNA mutations. Initially, it was hypothesized that each adduct resulted in a specific mutation.\(^8\) However, it became evident
that a single adduct could lead to a variety of mutations, depending on the base sequence context.\textsuperscript{9} Such context-sensitive variation involves the adoption of different conformations within DNA.\textsuperscript{9,10}

The nature of the mutation produced during replication is not only specific to the damage within the DNA, but is modulated by the DNA polymerases involved.\textsuperscript{11} High-fidelity DNA polymerases replicate unmodified DNA bases with high efficiency but are generally intolerant of the DNA distortions caused by many types of DNA lesions, and therefore, are blocked or misinsert bases opposite DNA adducts during replication.\textsuperscript{2} The active sites of these enzymes enclose the substrate DNA tightly and exclude solvent; this tight enclosure is initiated by the binding of the complementary dNTP to a zinc finger in the enzyme.\textsuperscript{12} The enzymes also establish multiple contacts to the minor groove, which is the basis of their proofreading function; any mismatches they detect are relayed to an exonuclease domain. In contrast, DNA polymerases of the Y family, which engage in “translesion synthesis,” contain larger active sites which mediate DNA replication across various replication stalling DNA lesions; however, even translesion polymerases can be inhibited by certain DNA lesions and bypass has been shown to be error-prone depending on the adduct.\textsuperscript{13} Y-Family polymerases have more spacious and solvent-accessible active sites, they lack exonuclease domains or contacts to the minor-groove edge of template-primer.\textsuperscript{14}

Adduct formation at dG has been extensively documented. Recall that dG has the lowest oxidation potential of the nucleobases; it is the most electron rich nucleoside and therefore reacts readily with various DNA damaging electrophiles. Modification of dG results in alteration of the \textit{anti} \textit{syn} glycosidic bond equilibrium, which gives rise to multiple conformations within the duplex; this can be attributed to the increased steric bulk of the adduct.\textsuperscript{15} The situation becomes more convoluted when one considers other factors that will influence this equilibrium within the duplex including adduct type, site of attachment, flanking bases, complementary base and environment (e.g. polymerase active site). Therefore, the conformational flexibility of bulky
DNA adducts may cause various pronounced perturbations of the helix and lead to mixtures of targeted (i.e. base-substitution) and semi-targeted (i.e. frame-shift) mutations if unrepaired.

Experimental studies have characterized numerous authentic DNA adducts obtained from various classes of carcinogens; the most attention has been paid to compounds for which there is substantial evidence of human exposure, or for which adducts have been shown to form in vivo. In general, adduct formation can result from N-nitrosamines,¹⁶ aflatoxins,¹⁷ polycyclic aromatic hydrocarbons (PAHs),¹⁸ estrogens,¹⁹ phenoxy radicals,²⁰ aromatic amines,²¹ and aryl hydrazines.²² Typically, in this field, chemists have followed toxicologists: they study the impact of a modification within DNA once authentic adducts have been identified in a biological system (or organism).

**N-Nitrosamines**

N-Nitrosamines are present in various foods, beverages, tobacco, cosmetics, cutting oils, hydraulic fluids and rubber products.²³ The metabolic activation of these chemically inert compounds into reactive electrophiles typically involves the α-hydroxylation by several cytochrome P450 enzymes.²³ Such metabolites are unstable and rapidly decompose to produce aldehydes and alkyl diazohydroxides, the latter of which have the ability to alkylate DNA. A general rule regarding these alkylating agents is that the potential exists to form adducts with all exocyclic oxygens and ring nitrogens, with the exception of the N1 site of guanine.²³

**Aflatoxins**

Humans are exposed to aflatoxins through the consumption of mouldy cereals, grains, and nuts.¹⁷ Aflatoxin B1 (AFB1) is the most abundant constituent of this class as well as the most carcinogenic. Metabolic activation of AFB1 involves oxidation of an alkene to an epoxide²⁴ which reacts with DNA to yield primarily dG adducts at N7.¹⁷
Polycyclic aromatic hydrocarbons (PAHs)

PAHs are by-products of combustion processes, resulting in ubiquitous human exposure to this class of carcinogens. They have been implicated in lung cancer due to their presence in tobacco smoke. The most common PAH compound, benzo[a]pyrene (B[a]P), has been classified by the International Agency for Research on Cancer (IARC) as a carcinogenic substance to humans. Carcinogenic PAHs share some common structural features. They generally have four or more fused benzene rings and include a bay region. When the bay region is either methylated or closed by another ring, a sterically hindered fjord region is created, which decreases the planarity of the aromatic system and increases the potency of the carcinogen. Metabolic activation of PAHs typically results in diol epoxides, which react with nucleophilic centers such as N2 of dG. Alternatively, C8-dG adducts arising from PAH exposure include C-bonded lesions generated via a radical cation intermediate along the metabolic pathway (Figure 3 - 1 (a)).

![Figure 3 - 1: Structure of various C8-aryl-dG analogues. (a) Adduct formation from PAH carbocation formation; (b) adduct formation due to estroquinone radical formation (c) adduct formation from activation of OTA; (d) adduct formation from phenoxyl radical, or aryl radical via reduction of arenediazonium ions.](image-url)
Estrogen Adducts

Human estrogens have been classified as carcinogens by the IARC\(^\text{29}\). One documented mechanism of action involved the activation of estrogens to o-quinone intermediates that form DNA adducts. In particular, the estrogen 3,4-estronequinone (3,4-EQ) undergoes metabolic activation to a 3,4-EQ radical anion, which generates C8-dG adducts such as 8-3,4-EQ-dG (Figure 3 - 1 (b)).\(^\text{30}\) Relatively little is currently known about these adducts and their connection to cancer is currently a topic of interest.\(^\text{31}\)

Phenoxyl Adducts

Several phenol-containing compounds have been shown to cause oxidative stress and DNA damage in cells and are classified as environmental carcinogens by IARC.\(^\text{20}\) Ochratoxin A (OTA) is a para-chloro-phenolic mycotoxin produced by several species of fungi.\(^\text{32}\) OTA has been shown to form a carbon-bonded C8-dG adduct following oxidative activation by redox-active transition metals (Figure 3 - 1 (c)).\(^\text{20}\) Phenoxyl radical formation can be mediated in humans by peroxidase enzymes or redox active transition metals.\(^\text{20}\) Studies have shown that unsubstituted C8-dG phenoxyl adducts are formed when phenoxyl radicals react with dG.\(^\text{20}\) Phenoxyl adducts have also been implicated in the genotoxic effects of benzene which can be metabolized to p-benzoquinone and hydroquinone.\(^\text{33}\) Previous work in the Manderville group has aimed at understanding the conformational preference of the ortho- and para-phenol dG lesions in DNA.\(^\text{34}\) Both phenol adducts destabilized the duplex when base paired with C but show a sequence-dependent increase in duplex stability when mismatched with G, which may lead to mutagenic hotspots.\(^\text{35}\)

Aromatic Amines (AAs)

Human exposure to aromatic amines and amides occurs from a number of sources, including various industrial processes, incomplete combustion, cigarette smoke and certain foods
(i.e. charred meats). The activation of aromatic amines and amides generally consists in an N-oxidation to yield N-hydroxy arylamines and N-hydroxy arylamides (arylhydroxamic acids). N-Hydroxy arylamines can react directly with DNA, while arylhydroxamic acids are not directly electrophilic and must be further metabolized to reactive esters. Typically, the major adducts that arise from these electrophilic intermediates are formed through covalent linkage of the amine or amide nitrogen to C8 of dG, whereas minor adducts arise from reactions with the exocyclic nitrogens and oxygens of dG and dA in vivo.

The most studied AAs include 2-aminofluorene (AF) and N2-acetylaminofluorene (AAF) (Figure 1 - 18), which were initially developed as pesticides, but were banned before use due to their carcinogenicity. This carcinogenicity is positively correlated to the formation of DNA adducts at both the N2 and C8 position of dG. As mentioned in Section 1.1.6.5, three conformational motifs have been identified for a variety of AA adducts (B, S, W conformations) which exist in equilibrium. The flexibility of the bulky group relative to the dG moiety also influences the conformer populations, such that greater co-planarity results in higher levels of the stacked S-conformation. Although AAF-dG and AF-dG have similar sized bulky groups, AF-dG is capable of adopting both the syn stacked (S) and anti (B) conformers. AAF-dG possesses relatively limited conformational flexibility compared with AF-dG, with predominant syn conformation, which is likely due to steric constraints imposed by the acetyl group at the N-linker site.
In order to understand how the structural differences between C8-bonded AF-dG and AAF-dG influence their ability to affect DNA replication, crystal structures of DNA polymerases complexed with AAF-dG or AF-dG modified DNA were analyzed.\textsuperscript{12,43-45} AF-dG bound to DNA polymerase adopts the \textit{anti} conformation, which led to WC base pairing with the incoming dCTP in the primer strand and permitted replicative bypass with high fidelity polymerases.\textsuperscript{43,44} In contrast, AAF-dG adopts the \textit{syn} conformation, which blocked replication by high fidelity polymerases and required low fidelity polymerases (e.g. Y-family polymerase η) to slowly bypass the lesion in an error-free manner.\textsuperscript{45} These studies provided a correlation between the structures of the adducts in the polymerase active site and those in duplex DNA.\textsuperscript{43}

Within DNA strands AF adducts primarily lead to base substitutions.\textsuperscript{46} Studies have revealed stabilization of a G:A mismatch\textsuperscript{47} where AF-dG obtains a \textit{syn} conformation, with Hoogsteen H-bonding to A in double stranded DNA.\textsuperscript{48} This provides a possible rationale for the formation of the G to T transversion mutation most frequently observed for AF-dG.\textsuperscript{47,49} In contrast, AAF-dG primarily causes two-base deletions which correlates with its large preference
for a syn conformation resulting in the AAF moiety being within the base stack, resulting in bulge stabilization.\textsuperscript{50}

In addition to the above studies on AF-dG and AAF-dG adducts, nucleotide excision repair (NER) efficiencies have been studied and compared with those related to PAH lesions.\textsuperscript{51} It was found that the lesion forms a base-displaced intercalated conformation in dsDNA and was resistant to NER when the nucleotide opposite the lesion was removed.\textsuperscript{52} The lesion was repaired with high efficiency in full complementary duplexes. However, the efficiency of excision is reduced by 15-20\% in deletion duplexes compared with the corresponding full length duplex. The results were explained by the relative stability of the modified duplex, which is an important factor for lesion recognition by NER enzymes.

**Aryl Hydrazine Adducts**

Humans are exposed to aryl hydrazines through industrial, pharmaceutical and dietary sources.\textsuperscript{53} Several aryl hydrazines are found in mushrooms (Agaricus bisporus). Aryl hydrazines with various para substituents have been shown to be metabolized to arenediazonium ions and then to aryl radicals that form carbon-bonded C8-aryl-dG adducts.\textsuperscript{22} These adducts seem to adopt unique structures, and their biological effects differ from those of AA lesions because they lack a heteroatom in the nucleobase-aryl linkage. C8-aryl hydrazine adduct formation decreases the stability of N-glycosidic bonds, which gives rise to abasic sites.\textsuperscript{22} One specific study incorporated \textsuperscript{10}Be-dG (Figure 3 - 1 (d)) into oligonucleotides that were subsequently used as templates for primer extension by various DNA polymerases, and it found that these lesions stalled DNA synthesis.\textsuperscript{54} Small amounts of dA and dG misincorporation, as well as one- and two-base deletions were observed. The anti/syn conformational heterogeneity of \textsuperscript{10}Be-dG is believed to be responsible for the mutagenic outcomes.\textsuperscript{54}
3.2.2. Determination of Structure-Activity Relationships

The correlation between DNA adduct structure and conformation on the one hand, and their biological consequences on the other, has been addressed experimentally by the creation of a variety of lesions, whose functional effects were then determined.

Studies by Zhang and coworkers have studied incorporation and extension past various N\(^2\)-guanine adducts (Figure 3 - 2).\(^2\) The N\(^2\) atom of guanine is susceptible to modification by various potential carcinogens, including formaldehyde,\(^55\) acetaldehyde,\(^56\) styrene oxide,\(^57\) N-nitrosopyrrolidine,\(^58\) and the oxidation products of various PAH’s.\(^27\)

Zhang and coworkers have altered the size of the adduct at N\(^2\) to determine its effect on incorporation and extension past the lesion.\(^2\) Their series included N\(^2\)-MeG, N\(^2\)-EtG, N\(^2\)-IbG, N\(^2\)-BzG, N\(^2\)-NaphG, and N\(^2\)-AnthG. They concluded that the polymerase efficiency and fidelity of incorporation opposite the lesion was largely preserved with increasing adduct size (i.e. the correct dC base was inserted opposite the lesion). All adducts except N\(^2\)-AnthG were readily extended beyond the modification by Dpo4 and yielded full-length products, although some unelongated product was isolated. N\(^2\)-AnthG was found to be a severe block for which only a one-base extension product was observed. Two X-ray structures were determined for Dpo4 in complex with duplex DNA, for which the template strand contained the N\(^2\)-NaphG residue. The duplex was designed such that a 14-base primer contained a dideoxy-cytosine at the terminus which paired with the N\(^2\)-NaphG template residue. The lack of a 3′OH resulted in the N\(^2\)-NaphG remaining in the post-insertion (-1) position. Interestingly, the N\(^2\)-NaphG residue was found to be in the syn orientation in one complex (termed complex 1) and anti orientation in the other (termed complex 2). The terminal C (dideoxy-C) of the primer in complex 2 was found to be flipped into the minor groove resulting in a non-productive Dpo4 structure. In both complexes there was considerable distortion of the DNA duplex when compared with the structures of DNA lesions that replicated with high efficiency.\(^2\) The base pairs to the 3′-side of the N\(^2\)-NaphG residue were
severely buckled, with large tilts and increased base pairing distances between hydrogen bond donors/acceptors several residues upstream from \( N^2 \)-NaphG, which indicated that the lesion can distort the helix several base pairs away from the actual site of damage.

The findings by Zhang and coworkers were related to a study in which a Benzo[a]pyrene-diol-epoxide-dG (BPDE-G) adduct was observed in two different orientations when bound to Dpo4.\(^\text{59}\) In this example, the \textit{syn} orientation of the \( N^2 \)-BPDE-G adduct in the Dpo4 active site remained non-productive, as the BPDE ring system stacked between the last two bases in the primer and prevented optimal binding of the incoming triphosphate. This contrasts with the \( N^2 \)-NaphG moiety in the \textit{syn} orientation, as Dpo4 can achieve what appeared to be a productive post-insertion complex. The difference between the adducts, aside from the size of the aryl moiety, is the extra flexibility of a methylene linker group, which allows for the \( N2 \)-NaphG adduct to obtain productive conformations.

Although these studies have provided some insight into the implications of adduct conformational heterogeneity, they have been focused on the \( N2 \) site of dG for modification. Clearly, modification at C8 will result in different perturbations within the duplex. Although modification at either site affects the \textit{anti/syn} conformational equilibrium, their mechanisms differ, as modification at \( N2 \) directly impacts the W-C H-bonding face of the nucleoside, while modification at C8 does not. Furthermore, as was seen when comparing the BPDE-G adduct and the \( N^2 \)-Naph-G adduct, additional flexibility between the dG base and the aryl moiety has implications for mutagenicity.

As previously mentioned, the most extensively studied class of carcinogens that react at the C8-position of guanine are the aromatic amines. There are two types of aromatic amines: non-acetylated lesions, which reduce the replication efficiency but are in general faithfully bypassed by high fidelity polymerases (small amounts of transversion mutations);\(^\text{60}\) and the
acetylated derivatives, which block replicative polymerases, but can be bypassed with special low-fidelity polymerases\textsuperscript{60} and have been shown to be mostly error-free,\textsuperscript{61} although frameshift mutations may occur in repetitive sequences.\textsuperscript{60}

Figure 3 - 3: (a) AF-dG lesions exist in an \textit{anti-syn} conformational equilibrium. (b) AAF-dG lesions exist primarily in the \textit{syn} orientation. (c) Structure of various C8-N-acetyl-aryl-dG lesions studied.

The distinct mutagenic properties of the acetylated and non-acetylated aromatic amine lesions result from their conformational preferences. The non-acetylated lesions exist in an
equilibrium of syn and anti conformations. The acetylated lesions adopt the syn-conformation with high preference, which is independent of the size of the aromatic unit. Mechanistic insights into the replication through acetylated AAF-dG lesions using low-fidelity polymerases was provided by Schorr and coworkers. They found that the polymerase correctly incorporated, with high specificity, a dC opposite an AAF-dG or AAA-dG lesion despite the presence of the aromatic unit attached to C8; full extension was also observed, although it proceeded significantly more slowly. The authors hypothesized that, due to the large preference of the acetylated lesions for the syn conformation, it was unlikely that these adducts would obtain anti structures to permit replicative bypass. Through the use of crystallography, they noted that for duplexes containing AAF-dG and AAA-dG adducts the primer strand rotated around the bulky moiety, which maintained the syn conformation of the adduct to allow bypass by a low fidelity polymerase. Furthermore, they found that bypass occurred less commonly with increasing size of the adduct and was entirely absent with pyrene adducts (AAP-dG). This was attributed to the finding that the aryl moiety of the AAF-dG and AAA-dG adducts stacked on top of the W-C base pair formed at the terminus of the primer/template duplex. The bulky aromatic units blocked the active site for an incoming dCTP. Interestingly, the enzyme was able to rotate the primer strand, which brought the dG part of the lesion closer into the active site. This rotation resulted in unstacking of the aromatic moiety and partial opening of the active site for the complementary dCTP to bind. The major energetic penalty of this rotational process was the required unstacking of the aromatic bulky adduct unit and the last formed W-C base pair between the primer and template strand. Schorr et al. elegantly designed a family of modifications to determine if increasing the size of the aromatic adduct resulted in slower translesion synthesis (TLS). For this experiment, they studied a benzyl (AAB-dG), naphthyl (AAN-dG) and pyrene (AAP-dG) unit (Figure 3 - 3) and found that the TLS efficiency was strongly reduced upon increasing aromatic adduct size.
A separate interesting aspect of this work was the elucidation of how the correct dCTP base was inserted opposite the lesion. The authors attributed the observed selectivity of base insertion to the presence of a Hoogsteen, H-bonding interaction between the C4-NH$_2$ group of the dCTP and the C6=O group of the lesion. The formed critical H-bond discriminated between dCTP and dTTP, while the purine triphosphates could not bind due to their larger size. They demonstrated the importance of this hydrogen bond by using an artificial triphosphate, which was identical to dC but lacked the C4-amino group and was not inserted opposite the lesion.

One would expect the flexibility of the adduct to have a rather large influence on the ability of the aryl-amine moiety to stack with the terminal base pairs. In the case of AA adducts, increasing the size of the aryl moiety attached at the C8-site of dG results in larger π-surfaces, this allows for greater stacking interactions to occur with neighbouring nucleosides. This directly contrasts with C8-aryl-dG adducts, as the lack of a flexible linker atom results in a biaryl system. In biaryl systems the rings twist out of plane due to steric interactions; larger ring sizes typically result in larger degrees of twist. Therefore, for C8-aryl-dG systems, larger aryl moieties do not necessarily correlate with more stable stacking interactions, as the resulting steric clash of the aryl ring with dG does not allow stacking interactions to be optimized. This can result in large differences in mutagenicity. As shown by Schorr, ring size did not influence the syn/anti conformation for the acetylated derivatives, and larger ring sizes were more able to stack with terminal base pairs, which made their displacement by a polymerase more difficult. C8-aryl-dG adducts anti/syn conformation are significantly affected by aryl ring size, and their ability to participate in stacking interactions within a polymerase active site will clearly be modulated by the lack of flexibility between C8 and the aryl moiety.

3.2.3. Fluorescent Nucleobase Analogues

Fluorescent nucleic acid base analogues have become extremely important in the fields of biology, biochemistry and biophysical chemistry, as well as in the field of DNA nanotechnology.
Such molecules can be introduced as fluorescent probes into a nucleic acid of interest close to the site under study with minimal perturbation of the system.

Fluorescent nucleobase analogues have also been studied extensively in the literature, which reports that factors such as hydrogen bonding,\textsuperscript{64} single- or double-stranded environment,\textsuperscript{65} and neighbouring base govern the fluorescence quantum yield of a base analogue inside the base stack.\textsuperscript{66} Generally, the fluorescence exhibited by the base analogues is quenched inside DNA, and normally the effect is most significant if surrounded by neighbouring purines.\textsuperscript{65,67} One example, 2-AP, has a high fluorescent quantum yield (0.68) as a free base in aqueous solution at physiological pH and effectively base-pairs with thymidine.\textsuperscript{68} Upon duplex hybridization, the fluorescence of 2-AP is significantly quenched. To a lesser degree, quenching is also observed upon incorporation of 2-AP into single stranded regions. This has prompted 2-AP to be used to study nucleic acid structures including duplex structures, quadruplex loops and to probe enzyme-induced effects on nucleic acid structure.\textsuperscript{69}

Pteridines are another example of highly fluorescent nucleobase analogues (Figure 3 - 4). Three examples include 3-methyl isoxanthpterin (3MI),\textsuperscript{70} 6-methylisoxanthopterin (6MI)\textsuperscript{71} and 4-amino-6-methyl-8-(2-deoxy- D-ribofuranosyl)-7(8H)-pteridone (6MAP),\textsuperscript{72} which have quantum yields of 0.88, 0.70 and 0.39 as monomers, respectively. 6MI is a guanine analog that perfectly base pairs with cytosine.

![Figure 3 - 4: Structure of various pteridine fluorescent nucleobase analogues.](image-url)
The goal for these modified fluorescent nucleobases is to probe various DNA environments while retaining the native behaviour of the nucleic acid system being studied. Therefore, it is considered detrimental for probe incorporation to result in a reduction in thermal stability of the duplex, changing the H-bonding interactions, or large degrees of steric bulk.\(^7\)

**3.2.4. C8-Aryl-dG Modifications as “Probing Adducts”**

Compared to classical fluorescent nucleobase analogues, 8-aryl-dG adducts are used with a different purpose. They are not intended to represent the native behaviour of the oligonucleotide, but instead as models to study the effects of the mutagenic adducts on DNA structure. The adduct is the probe, and it provides the diagnostic handle to determine how various structural analogues impact thermal stability, how their steric bulk affects the nucleic acid structure, and how they are eventually processed by polymerases. These probing adducts are rooted in toxicology as they are structural analogues of authentic DNA adducts that have been shown to form from numerous sources as discussed (Figure 3 - 5). Their ability to function as structural analogues stems from their similarities in the attachment of the aryl system to the C8 site of dG, and the associated differences in aryl ring size and shape. Attachment of an aryl moiety at the C8 site of dG extends the purine π-system, which results in fluorescent nucleobase analogues that exhibit longer excitation and emission wavelengths than unmodified nucleobases, allowing them to be selectively excited in the presence of nucleic acids and proteins. The Manderville group has previously published on the use of various C8-aryl-dG adducts to probe conformation within duplex DNA,\(^34,35,74,75\) metallation,\(^76\) and H-bonding.\(^77\) The ability of C8-aryl modified dG nucleosides to function as fluorescent probing adducts is unique; the presence of a flexible linker within the N-linked derivatives limits the conjugation of the π-system and they do not exhibit fluorescence.

The newly optimized solid phase DNA synthetic methodologies presented in Chapter 2 allowed us to prepare these adducted oligonucleotides in larger amounts and to employ them in
studies on mutagenicity. Furthermore, the generation of the authentic 22-NarI oligonucleotides has allowed for various primer extension assays to be completed by Anne Verway at the University of Guelph. These studies are ongoing.

Figure 3 - 5: Comparison of how the family of 8-aryl-dG probing adducts (bottom) compare as structural analogues of various biologically relevent DNA adducts (top).

Clearly, the interplay between the size of the bulky group, adduct planarity, sequence and interstrand hydrogen bonding has major implications in determining adduct conformation, duplex structure, and the associated mutagenicity. This has been extensively demonstrated for the aromatic amine adducts formed, and the literature is replete with studies regarding this family of adducts. However, direct relationships between the AA adducts mutagenicity cannot be extended to carbon linked C8-dG adducts due to the presence of the N-linker bonds in the AA adducts, which significantly alters their ability to attain planar, stacked structures. The various mNarI oligonucleotides containing C8-aryl-dG adducts synthesized in Chapter 2 are further characterized in this chapter to provide a systematic analysis and interpretation of how structural differences (as their aryl moieties are all distinct in terms of size, shape, and electronics) impact duplex structure. Although these adducts are not themselves found to form in vivo, they provide
structural insights into the specific interactions that occur when C-linked C8-aryl-dG adducts are present within the duplex environment. Furthermore, C8-aryl-dG adducts double as fluorescent probes, which allows them to literally shed light on the whole situation.

3.3. Materials and Methods

3.3.1. General Methods

A detailed outline of the general methods for experiments relating to Chapter 3 can be found in Appendix A.

3.3.2. Preparation of Oligonucleotide Substrates

Preparation of the fully protected PhdG, FurdG, BtdG, QdG, CNPhdG and PydG phosphoramidites were performed as outlined previously in Scheme 2 - 5. To study the impact of the 8-aryl-dG adduct size, shape and electronics on duplex structure, and to determine the probing performance within different duplex mismatch scenarios, they were incorporated into the 12mer NarI sequence (5′-CTCGGCXCCATC where X = modified base) as described in Appendix A. Oligonucleotide purification was completed as described within Appendix A. Following HPLC purification, the purified oligonucleotide was lyophilized to dryness and analyzed by ESI-MS as described in Appendix A. Mass analysis was presented in Table 2 - 1. Representative mass spectra can be found in Appendix C. Oligonucleotides were quantified according to procedures in Appendix A. Duplexes were annealed from equal amounts (6μM) of NarI(X) and NarI(Y) (complementary strands) for study, as outlined in Appendix A. These oligonucleotide sequences were as follows: NarI(C) (5′-GATGGCGCCGAG); NarI(G) (5′-GATGGGGCCGAG); NarI(10) (5′-GATGGCCGAG) and NarI(THF) (5′-GATGG(THF)GCCGAG), where THF is the synthetic equivalent of an abasic site.
3.3.3. Computer Simulations

All computer simulations and calculations were carried out by Purshotam Sharma, in the Wetmore group at the University of Lethbridge (Lethbridge, AB). This data is presented to aid in the description of the experimental results involving C8-aryl-dG adducts influence on duplex structures. Experimental details can also be found in Appendix A, which were provided by Purshotam Sharma.

3.4. Results and Discussion

3.4.1. Nucleoside Properties

Numerous C8-aryl-dG analogues have previously been reported within the Manderville group showing that the fluorescent base analogues have quantum yields that are somewhat to highly sensitive to their immediate surroundings.\textsuperscript{35,75-77} The solvatochromic properties of the C8-aryl–dG nucleoside adducts are given in Table 3 - 1.
Table 3 - 1: Solvatochromic properties of 8-aryl-dG adducts

<table>
<thead>
<tr>
<th>Property</th>
<th>Solvent (ε)</th>
<th>Adduct</th>
<th>Ph dG</th>
<th>Fur dG</th>
<th>Bth dG</th>
<th>Py dG</th>
<th>Q dG</th>
<th>CNPh dG</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ&lt;sub&gt;max&lt;/sub&gt; (nm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O&lt;sup&gt;a&lt;/sup&gt;</td>
<td>277, 4.33</td>
<td>292</td>
<td>315</td>
<td>342</td>
<td>313</td>
<td>308</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH&lt;sub&gt;3&lt;/sub&gt;CN</td>
<td>292</td>
<td>292</td>
<td>321</td>
<td>345</td>
<td>315</td>
<td>327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHCl&lt;sub&gt;3&lt;/sub&gt;</td>
<td>289</td>
<td>298</td>
<td>318</td>
<td>345</td>
<td>318</td>
<td>325</td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ&lt;sub&gt;em&lt;/sub&gt; (nm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O&lt;sup&gt;a&lt;/sup&gt;</td>
<td>395</td>
<td>384</td>
<td>419</td>
<td>(388)</td>
<td>407</td>
<td>468</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH&lt;sub&gt;3&lt;/sub&gt;CN</td>
<td>384</td>
<td>371</td>
<td>413</td>
<td>476.7</td>
<td>(384)</td>
<td>454</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHCl&lt;sub&gt;3&lt;/sub&gt;</td>
<td>373</td>
<td>366</td>
<td>405</td>
<td>431</td>
<td>468</td>
<td>424</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δν (cm&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10785</td>
<td>8205</td>
<td>7880</td>
<td>(3466)</td>
<td>8621</td>
<td>7379</td>
<td>11100</td>
<td></td>
</tr>
<tr>
<td>CH&lt;sub&gt;3&lt;/sub&gt;CN</td>
<td>8205</td>
<td>7592</td>
<td>6940</td>
<td>8021</td>
<td>12138</td>
<td>8554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHCl&lt;sub&gt;3&lt;/sub&gt;</td>
<td>7792</td>
<td>6235</td>
<td>6755</td>
<td>5783</td>
<td>10079</td>
<td>7184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Φ&lt;sub&gt;fl&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H&lt;sub&gt;2&lt;/sub&gt;O&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.44</td>
<td>0.49</td>
<td>0.46</td>
<td>0.08</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH&lt;sub&gt;3&lt;/sub&gt;CN</td>
<td>0.49</td>
<td>0.183</td>
<td>0.244</td>
<td>0.27</td>
<td>0.051</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHCl&lt;sub&gt;3&lt;/sub&gt;</td>
<td>0.22</td>
<td>0.025</td>
<td>0.394</td>
<td>0.29</td>
<td>0.185</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> 10 mM MOPS buffer, pH 7.0, m = 0.1 M NaCl. <sup>b</sup> Compound with donor acceptor character, log ε for high energy absorbance (values in brackets are for highest intensity peak). <sup>c</sup> Stokes shift (Δν) is calculated as (1/λ<sub>abs</sub> − 1/λ<sub>em</sub>). <sup>d</sup> Fluorescence quantum yields are relative to quinine bisulfate (Φ<sub>fl</sub> = 0.546 in 0.5 M H<sub>2</sub>SO<sub>4</sub>).

All adducts show increased Stokes Shifts with increased solvent polarity. This is consistent with a larger dipole moment in the excited state than the ground state. Ph dG, Fur dG and Bth dG adducts typically show reduced quantum yields as solvent polarity decreases; however, the Bth dG is quite emissive in CHCl<sub>3</sub>. These analogues do not exhibit donor-acceptor (D-A) character and are classified as “non push-pull analogues” which will be discussed further (vide
The PydG, QdG and CNPhdG adducts exhibit D-A character and constitute the “push-pull” family of adducts. Interestingly, the push-pull derivatives have quenched fluorescence in the presence of water. This has been observed with cyano-anilines and other push pull systems and has been attributed to an increase in the non-radiative rate of decay.\textsuperscript{78}

The CNPhdG adduct was hypothesized to model the classical donor acceptor (D-A) dimethylaminobenzonitrile (DMABN) moiety within DNA. The dual fluorescence (i.e. fluorescence emission occurring at two distinct wavelengths) of DMABN has been studied extensively, and numerous explanations for the dual fluorescence have been proposed.\textsuperscript{79,80} It now seems generally accepted that the dual fluorescence can be explained by a charge transfer (CT) model based on the benzene anion radical.\textsuperscript{81} Excitation of molecules with D-A character results in CT processes that transfer an electron from the HOMO of the donor to the LUMO of the acceptor. This process is depicted in Figure 3 - 6 for CNPhdG. CT generates charge separated species, and CT emission can occur from twisted (D*-A*) or planar geometries (D+=A-).

Figure 3 - 6: Generation of the CT species upon excitation of CNPhdG.
Upon excitation of DMABN, three low lying electronic excited states are expected, two form a CT state (which are either twisted or planar), and one a locally excited (LE) state. The two CT states have been classified as antiquinoid (AQ) or quinoid (Q) due to the Jahn-Teller induced distortion of the benzene anion radical. The dual fluorescence may arise from any two states; the CT states have a red shifted emission compared to the LE state, and each excited state has a different geometry at which it attains a minimum energy. The reaction coordinate involves charge transfer from donor to acceptor, intramolecular twisting between these subunits, and solvent relaxation around the newly created large dipole (which is largest when $\theta = 90^\circ$). Therefore, as one would expect, more polar solvents stabilize the larger dipole and can result in dual emission.

Unfortunately, CNPh-dG failed to produce dual fluorescence in CH$_3$CN, CHCl$_3$ or H$_2$O; however, it did show dual emission in DMSO at 371 and 470 nm (work by Katie Rankin not shown). Interestingly, the $^0$dG adduct did display dual fluorescence in CH$_3$CN (384 and 510 nm) while showing only LE emission in H$_2$O (407 nm) and only CT emission in CHCl$_3$ (468 nm) (Figure 3 - 7). As previously noted, H$_2$O (and other polar protic solvents) effectively quench the fluorescence for D-A, push-pull systems through increasing the rate of non-radiative decay via H-bonding. A possible depiction of the overall processes occurring is provided in Figure 3 - 7, which has been adapted from models proposed by Cogan et al. for their substituted benzene derivatives. Herein, excitation of the $^0$dG to its initially excited LE state equilibrates to an excited CT state, which is stabilized in polar solvents. This state is analogous to the CT (Q) state described by Cogan which was said to attain a perpendicular geometry with regards to donor and acceptor (i.e. $\theta = 90^\circ$). This perpendicular geometry generates a large dipole moment which stabilized by polar solvents. In the case of $^0$dG, CT state 1 is not emissive in water, as it undergoes radiationless decay to the ground state. Therefore, a small amount of LE emission is observed, which is attributed to the relatively small population of excited state $^0$dG that did not
equilibrate to the lower energy CT state and emitted as LE. In the polar aprotic CH$_3$CN (Figure 3 - 7 (B)), the CT emission is not quenched (nor is the perpendicular geometry as stabilized) as it was in water; however, the perpendicular CT (Q) state is somewhat stabilized in the polar CH$_3$CN solvent, and dual emission is observed from LE and CT.

Figure 3 - 7: The dual fluorescence observed for the push-pull $^0$dG adduct. (a) Observed emission spectra for $^0$dG in water (red trace), CH$_3$CN (dashed blue trace) and CHCl$_3$ (dotted green trace); (b) proposed emission in H$_2$O; (c) proposed emission in CH$_3$CN; (d) proposed emission in CHCl$_3$.

When $^0$dG was excited in CHCl$_3$ it produced an unexpected result; the adduct displayed a significantly red shifted emission (468 nm), consistent with a CT state (denoted CT state 2 in Figure 3 - 7 (C)). This state, analogous to the CT (AQ) state described by Cogan et al. was reported to attain a planar geometry ($\theta \approx 0^\circ$). The large dipole moment generated by a perpendicular geometry cannot be stabilized by CHCl$_3$, which is non-polar. This emission suggests that CT state 2 is lower in energy than either CT state 1 or the LE state in CHCl$_3$. This
is supported by a statement made by Cogan et al. that “in many cases the planar conformation of CT excited states is lower in energy than that of the LE state, and dual fluorescence can be observed also from planar structures.” Therefore, it seems reasonable to attribute the emission at 468 nm to arise from a CT state, and is being rationalized to arise from a planar geometry.

Insight into the structural features and electronic properties of these C₈-aryl-dG adducts was obtained from DFT calculations, carried out by the Wetmore laboratory at the University of Lethbridge, Lethbridge, Alberta. B₉dG and Fur-dG adduct calculations predicted twisted S₀ (B₉dG Θ = 66.8°, Fur-dG Θ = 50.1°) structures that became almost planar in S₁ (B₉dG Θ = 188.1°, Fur-dG Θ = 0.16°), which corresponds with the behaviour of biphenyl systems. Typically, a higher degree of twist causes a blue shift in the absorption spectrum due to the disruption of conjugation between the nucleobase and the C₈-aryl substituent. This was found to be the case in work by Dr. Katie Rankin who compared the various classes of C₈-heteroaryl-dG nucleosides (comparing 5 membered series and benzoheterocyclic series). For the D-A nucleosides presented, the QdG adduct is more twisted in the ground state than the CNPh-dG owing to its larger fused ring system. The absorption spectrum is in agreement with the larger degree of twist, as the QdG absorbance is blue shifted compared to CNPh-dG. Upon excitation, calculations show both CNPh-dG and QdG become more planar in S₁, although not to the extent of the non-push pull derivatives (which approach planarity). Although the Ph-dG, Fur-dG and B₉dG adducts are more planar in their S₁ state, the push pull derivatives exhibit significantly red shifted emission compared to the non-push-pull derivatives due to their CT character as discussed earlier.

From interpretation of the HOMO and LUMO structures (Figure 3 - 8) it is evident that the non push pull derivatives contain delocalized π orbitals with almost equal density on the nucleobase and aryl ring system. In the push pull derivatives the density of the π-orbitals in the HOMO is also evenly distributed for both CNPh-dG and QdG. Interestingly, the LUMO clearly
shows significant π-orbital density on the aryl ring system, consistent with the CT character in these systems.

<table>
<thead>
<tr>
<th>Adduct</th>
<th>Characteristic</th>
<th>Value</th>
<th>Ground State Molecular Structure</th>
<th>Orbital Diagram HOMO</th>
<th>Orbital Diagram LUMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhdG</td>
<td>$\chi (S_0)/\chi (S_1)$ (Deg)</td>
<td>66.8/72.4</td>
<td><img src="image1" alt="Orbital Diagram" /></td>
<td><img src="image2" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Theta(S_0)/\Theta (S_1)$ (Deg)</td>
<td>41.2/17.2</td>
<td><img src="image3" alt="Orbital Diagram" /></td>
<td><img src="image4" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu(S_0)/\mu(S_1)$ (Debye)</td>
<td>4.83/4.87</td>
<td><img src="image5" alt="Orbital Diagram" /></td>
<td><img src="image6" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td>BinG</td>
<td>$\chi (S_0)/\chi (S_1)$ (Deg)</td>
<td>66.8/69.7</td>
<td><img src="image7" alt="Orbital Diagram" /></td>
<td><img src="image8" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Theta(S_0)/\Theta (S_1)$ (Deg)</td>
<td>211.1/188.1</td>
<td><img src="image9" alt="Orbital Diagram" /></td>
<td><img src="image10" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu(S_0)/\mu(S_1)$ (Debye)</td>
<td>5.45/6.41</td>
<td><img src="image11" alt="Orbital Diagram" /></td>
<td><img src="image12" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td>ForG</td>
<td>$\chi (S_0)/\chi (S_1)$ (Deg)</td>
<td>50.1/48.5</td>
<td><img src="image13" alt="Orbital Diagram" /></td>
<td><img src="image14" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Theta(S_0)/\Theta (S_1)$ (Deg)</td>
<td>343.8/0.16</td>
<td><img src="image15" alt="Orbital Diagram" /></td>
<td><img src="image16" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu(S_0)/\mu(S_1)$ (Debye)</td>
<td>4.27/4.86</td>
<td><img src="image17" alt="Orbital Diagram" /></td>
<td><img src="image18" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td>C14NHdG</td>
<td>$\chi (S_0)/\chi (S_1)$ (Deg)</td>
<td>68.4/44.0</td>
<td><img src="image19" alt="Orbital Diagram" /></td>
<td><img src="image20" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Theta(S_0)/\Theta (S_1)$ (Deg)</td>
<td>37.9/89.7</td>
<td><img src="image21" alt="Orbital Diagram" /></td>
<td><img src="image22" alt="Orbital Diagram" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu(S_0)/\mu(S_1)$ (Debye)</td>
<td>8.94/8.02</td>
<td><img src="image23" alt="Orbital Diagram" /></td>
<td><img src="image24" alt="Orbital Diagram" /></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3 - 8: Twist angles, dipole moments and ground-state B3LYP/6-31G(d) global minima orbital density plots of 8-aryl-dG adducts. $\chi$ refers to the glycosidic bond angle, $\Theta$ refers to the twist angle between the C8 site of dG and the aryl ring, and $\mu$ refers to the dipole moment of the nucleoside adduct. $S_0$ and $S_1$ refer to ground and excited electronic energy states respectively.
3.4.2. 8-Aryl-dG Adduct Impact on dsDNA Structure and Stability

3.4.2.1. Thermal Melting Analysis

Thermal Melting experiments were conducted in order to assess the impact of 8-aryl-dG modifications on the overall stability of the NarI duplex. Various conformations 8-aryl-dG adducts can attain within B-duplexes were presented in (Figure 1 - 19). Duplexes were prepared to a 6 μM concentration with various complementary sequences, as can be seen below (Scheme 3 - 1). NarI(G) was selected to probe the ability of the adduct to stabilize a mismatched pair. Recall that mismatch stabilization can lead to base transversions for the N-linked class of adducts.83 NarI(10mer) was selected to simulate the formation of a slippage product, which results in a bulge in the duplex. Stabilization of the bulge, resulting from the slipped base pairing, would indicate the ability of the adduct to induce a frameshift mutation and also have implications in NER as was shown with various PAH and AA adducts.83 Selection of NarI(THF) was used to probe the ability of the modification to stabilize a base displaced S structure by the adduct. With the lack of W-C base pairs, any stabilization of the abasic site duplex, relative to the unmodified duplex, would be attributed to favourable stacking of the adducted moiety within the DNA helix.
Scheme 3 - 1: Thermal melting analysis of 8-aryl-dG adducts within various NarI duplexes. (a) NarI(X):NarI'(C); (b) NarI(X):NarI'(G) (c) NarI(X):NarI'(10mer) NarI(X):NarI'(THF). Dotted black line X = dG, Red line X = Fur dG, Green line X = Bth dG, Blue line X = Q dG, Orange Line X = Py dG, Purple line X = CNPh dG, solid black line X = Ph dG.

<table>
<thead>
<tr>
<th>NarI X</th>
<th>NarI’ Strand</th>
<th>Na⁺ Duplex</th>
<th>Tm (°C)</th>
<th>ΔTm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dG</td>
<td>C</td>
<td>63.6 ± 1.1</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>54.1 ± 1.9</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>39.4 ± 1.1</td>
<td>/</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>45.7 ± 1.2</td>
<td>/</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>CNPh dG</td>
<td>C</td>
<td>48.9 ± 1.1</td>
<td>-14.6</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>55.5 ± 1.3</td>
<td>1.5</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>38.8 ± 1.4</td>
<td>-0.6</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>47.3 ± 1.3</td>
<td>1.6</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>Ph dG</td>
<td>C</td>
<td>54.7 ± 1.7</td>
<td>-15.7</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>47.8 ± 1.5</td>
<td>/</td>
<td>-0.6</td>
</tr>
<tr>
<td>10mer</td>
<td>36.5 ± 1.3</td>
<td>-2.9</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>42.2 ± 1.5</td>
<td>-3.5</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>Fur dG</td>
<td>C</td>
<td>54.5 ± 1.3</td>
<td>-9.1</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>57.6 ± 1.3</td>
<td>3.6</td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>36.8 ± 1.2</td>
<td>-2.6</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>42.7 ± 1.4</td>
<td>-3</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>O dG</td>
<td>C</td>
<td>44.99 ± 1.2</td>
<td>-18.6</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>45.0 ± 0.9</td>
<td>-9</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>39.4 ± 1.1</td>
<td>0</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>39.9 ± 1.7</td>
<td>-5.8</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>Bth dG</td>
<td>C</td>
<td>47.8 ± 1.6</td>
<td>-15.8</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>C</td>
<td>54.9 ± 1.5</td>
<td>0.9</td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>43.2 ± 1.7</td>
<td>3.8</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>50.1 ± 1.4</td>
<td>4.4</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>Py dG</td>
<td>C</td>
<td>43.5 ± 1.3</td>
<td>-20.1</td>
<td>/</td>
</tr>
<tr>
<td>G</td>
<td>49.0 ± 1.1</td>
<td>-5</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>10mer</td>
<td>45.6 ± 1.2</td>
<td>6.2</td>
<td></td>
<td>/</td>
</tr>
<tr>
<td>THF</td>
<td>45.3 ± 1.6</td>
<td>-0.4</td>
<td></td>
<td>/</td>
</tr>
</tbody>
</table>
Analysis of the $T_m$ data shows universal destabilization ($\Delta T_m = -9.1$ to $-20.1$ °C) of all NarI(X):NarI'(C) duplexes. The magnitude of destabilization increases in the order of $\text{Fur}dG < \text{CNPh}dG < \text{Ph}dG < \text{Bth}dG < \text{Q}dG < \text{Py}dG$. Recall that for N-linked C8-dG adducts base-paired with C, an equilibrium exists between the B- and S-type conformational motifs, with an exchange time in the millisecond range. Both of these conformations are destabilizing to the duplex. The duplex destabilization observed for C8-aryl-dG adducts paired with dC suggests that these adducts possibly exist in a B- or S-type duplex structural motif, with an anti or syn conformational preference, although there is not definitive evidence to distinguish between the two from $T_m$ data alone.

In a B-type motif, it is expected that the C8-aryl-dG adducts $\text{Fur}dG$, $\text{Bth}dG$, $\text{Q}dG$ and $\text{Ph}dG$, $\text{CNPh}dG$ and $\text{Py}dG$ would adopt the anti conformation within NarI upon base-pairing to their natural pyrimidine partner, placing the C8-aryl moiety in the major groove. Recall that the anti conformation is not energetically favoured for C8-aryl-dG adducts. Adoption of a B-type conformation could be a contributing factor to the destabilization of C8-aryl-dG modified duplexes. Furthermore, the location of the C8-aryl moiety in the major groove results in exposure of the lipophilic C8-aryl moiety to the aqueous extra-helical environment, resulting in duplex destabilization. The degree of duplex destabilization can be correlated to the relative polarity of the adducts. The ground state dipole moments ($\mu(S_o)$) of the C8-heteroaryl-dG adducts, as determined by DFT calculations, were presented in Figure 3 - 8. The adducts’ polarity decreases in the order of $\text{CNPh}dG > \text{Bth}dG > \text{Ph}dG > \text{Q}dG \approx \text{Fur}dG$ (calculations were not performed on $\text{Py}dG$).

One would expect that the adducts that display the largest ground state dipole moment are the most effectively stabilized by a polar environment as compared to ones with a lower ground state dipole moment. The most direct example showing stabilization for more highly polarized adducts in aqueous environments is seen by comparing $\Delta T_m$ values for $\text{CNPh}dG$ and $\text{Ph}dG$, as $\text{CNPh}dG$ is less destabilizing than the $\text{Ph}dG$, within NarI(X):NarI'(C). Additionally, another major contribution to
duplex destabilization can result from steric clash between the substituent located in the major groove and phosphate backbone and sugar moieties. This is observed within the $\Delta T_m$ trend, as the values correlate with size of the adduct as the bulkiest modifications are the most destabilizing ($\text{Fur}dG < \text{CNPh}dG < \text{Ph}dG < \text{Bth}dG < \text{Q}dG < \text{Py}dG$).

By contrast, duplex destabilization could be the result of the C8-aryl-dG adduct adopting a syn conformation in an S-type duplex motif, when base-paired with C. In this structure, the C8-heteroaryl moiety is intercalated between its flanking nucleobases, in turn flipping the opposing pyrimidine out of the helix. The resulting helical distortion and loss of W-C H-bonding would account for the decreased duplex stability, and decrease in $T_m$, compared to the unmodified duplex. The observed B/S-type heterogeneity in bulky N-linked C8-arylamine-dG adducts base-paired with C is known to be sequence-dependent, with the flanking bases of the adduct influencing the preferred conformation. For the AF-dG adduct, the abundance of the S-type conformation was obtained by the Cho group for a central trimer portion of a 12-mer duplex using $^{19}$F NMR, in which the modified base ($X = \text{AF-dG}$) was flanked by purines, pyrimidines or one of each type, in differing sequences. The Manderville group also recently demonstrated the sequence dependency of B- versus S-type in 10-mer $^{\text{Bth}}dG$-modified duplexes. In a duplex with the adduct flanked on the 5'-side by T and on the 3'-side by C, $^{\text{Bth}}dG$ was determined to predominantly adopt a syn conformation and S-type structure. On the other hand, in a duplex with the adduct flanked on the 5'-side by G and on the 3'-side by A, $^{\text{Bth}}dG$ was determined to adopt an anti conformation and B-type structure. These conformations were confirmed by both experimental CD and MD analysis, and were also supported by the % S-type data from Cho and co-workers. Of particular interest to the study presented here is the % S-type value determined by the Cho group for a 5'-CXC trimer, as the C8-aryl-dG adduct is also flanked on both the 5' and 3' sides by C in the NarI sequence. The 5'-CXC trimer was calculated to exist as 47 % S-type
conformer,\textsuperscript{86} suggesting that the \textit{anti} conformation, with B-type structure, is slightly more preferred for \textit{NarI} incorporated C8-heteroaryl-dG adducts base-paired with C.

Interestingly, when the modified \textit{NarI} oligonucleotides were hybridized to the \textit{NarI}'(G) complement, thereby introducing a mismatch, C8-aryl-dG adducts had variable influences on duplex stability. Stabilization occurred for the $^{Fwd}$dG (+3.6°C), $^{CNPb}$dG (+1.5°C), $^{BTh}$dG (+0.9°C) and $^{Pb}$dG (+0.7°C) adducts, while the $^{Q}$dG (-9°C) and $^{Py}$dG (-5°C) were destabilizing compared to that of the unmodified duplex. This stabilization may be due to the 8-aryl-dG adducts adopting a \textit{syn}-conformer, with a W-type duplex structural preference.\textsuperscript{74} Recall that the \textit{syn} conformation is energetically favoured for C8-aryl-dG adducts, and thus in turn, adoption of this conformation could be a contributing factor to the stabilization of C8-aryl-dG modified duplexes. Furthermore, in this conformation, the lipophilic C8-aryl moiety is located in the minor groove in a less solvent-exposed environment, and is thus favourably protected from the aqueous environment surrounding the duplex. Placing the C8-substituent in the minor groove typically enhances π-stacking interactions with neighbouring nucleobases; base-pairing stability is also significantly impacted when the adduct adopts a \textit{syn} conformation, as a result of Hoogsteen H-bonding of the G component of the C8-modified base with the W-C face of the opposing G.\textsuperscript{35,75} This stabilization effect was previously observed to occur with both phenolic\textsuperscript{35} and benzothienyl\textsuperscript{75} C8-dG modified 10-mers hybridized to the complement strand with dG opposing the adduct. These C-linked phenolic and benzothienyl C8-dG modified duplexes, in addition to N-linked C8-arylamine-dG modified duplexes,\textsuperscript{87} have been shown to favor a W-type motif, with the adduct in the \textit{syn}-conformation, when the adduct is mismatched with a purine.

The $T_m$ data seem to suggest that stabilization is shape dependent as one can see that the single rings and “linearly” shaped $^{Bb}$dG adduct were stabilizing, while the more “horizontally” shaped $^{Q}$dG and $^{Py}$dG were destabilizing. Due to the large aryl ring sizes of $^{Q}$dG and $^{Py}$dG, it is possible that, while these aryl moieties may still be located in the minor groove, they are not
ideally located to derive enhanced stability from additional π-stacking and base-pairing interactions. The data does not necessarily correlate to the “bulkiness” of the adduct as PydG is less destabilizing to duplex structure than the QdG. This can be justified based on the PydG’s increased “linear” component (i.e. benzene rings projecting further from the dG base) for which more optimal stacking interactions can occur with the flanking bases.

The NarI(10mer) data provide insight into stabilization of a slipped-mutagenic duplex. The Rizzo group reported that hybridization of an N-linked-dG adduct within the NarI sequence to the complementary 10-mer resulted in an increased $T_m$ of +10 °C, as compared to the unmodified duplex. It should be noted that the modification was at the same position in the NarI sequence as the C8-aryl-dG modifications studied here. In addition, the Lefèvre group incorporated the AAF-dG adduct into a central site within a GC-rich 12-mer, and found that upon hybridization to a complementary 10-mer, the duplex was stabilized, with an increase in $T_m$ of +15 °C.

Of the C8-aryl-dG adducts studied here, the single ring modifications mildly destabilized the duplex structure, while the larger fused ring systems either stabilized it or behaved neutrally. The trend in observed $\Delta T_m$ values increases in the order of $\text{PydG} (-2.9\, ^\circ\text{C}) \approx \text{Fur}dG (-2.6\, ^\circ\text{C}) < \text{CNPh}dG (-0.6\, ^\circ\text{C}) < QdG (0\, ^\circ\text{C}) < \text{Bth}dG (+3.8\, ^\circ\text{C}) < \text{PydG}(+6.2\, ^\circ\text{C})$ which correlates with the size and shape of the adducts. These findings suggest that a larger, longer aryl ring is required to stabilize the bulge formed by duplex slippage. The bulge in the duplex created by the deleted nucleotides in the 10-mer complement provides ample space for intercalation of the C8-aryl moiety into the minor groove. It is thought that smaller C8-aryl rings are not able to stabilize the bulge, as the small π-surface area would not be able to effectively participate in π-stacking with the flanking nucleobases. The large fused aryl ring size moieties are able to favourably interact with the π-stack, thereby increasing overall duplex stability. Importantly, while the large size of these fused rings promotes mismatch formation and thus destabilizes the helix, it also allows for
stronger $\pi$-stacking interactions within the bulge of the slippage product. The role of molecular shape in the stabilization of the fused ring systems is again reinforced, as the $^{\text{Bth}}$dG and $^{\text{Py}}$dG adducts are the most stabilizing; their “length” allows for the most optimized stacking interactions.

The $\text{NarI}^\prime$(THF) duplex was studied because it does not contain a complementary nucleobase that could engage in W-C or Hoogsteen H-bonding, thus removing these potential effects and allowing for sterics, $\pi$-stacking and solvation interactions to be probed more specifically. It is expected that in this scenario the unmodified dG base present in the $\text{NarI}$ strand is in an anti glycosidic conformation, as this allows for the largest degree of intra-strand $\pi$-stacking and retention of the natural B-DNA local structure.

The resulting $T_m$ values of these duplexes were specific to the 8-aryl-dG adduct with most modifications being destabilizing. Within the series of adducts the $\Delta T_m$ values ranged from $^{\text{Q}}$dG (-5.8°C) $< ^{\text{Ph}}$dG (-3.5°C) $< ^{\text{Fur}}$dG (-3°C) $< ^{\text{Py}}$dG (-0.4°C) $< ^{\text{CNPh}}$dG (+1.6°C) $< ^{\text{Bth}}$dG (+4.4°C). These results turned out to be quite informative. Unlike the $\text{NarI}^\prime$(10mer), which had extra space around the adduct, the $\text{NarI}^\prime$(THF) has geometrical constraints similar to the $\text{NarI}^\prime$(C) and $\text{NarI}^\prime$(G) duplexes with the complementary base flipped out. In a $\text{NarI:} \text{NarI}^\prime$(C) scenario, the ability of an 8-aryl-dG modification to displace the opposing base depends on several factors. Orienting the 8-aryl-dG nucleoside so that the aryl moiety is within the base stack should result in some stabilization due to increased $\pi$-stacking, and from sequestering the hydrophobic aryl moiety from an aqueous environment. These stabilizing influences are opposed by the energetic penalties resulting from flipping both bases from anti to syn, which results in the loss of W-C or Hoogsteen H-bonding, and unfavourable steric interactions. For all adducts except $^{\text{Bth}}$dG, $\pi$-stacking contributions seem unable to overcome such energy penalties. The $^{\text{Bth}}$dG, being the “longest and slimmest” adduct, seems the best able to stack within the core of the duplex, which provides the best opportunity for increased $T_m$ values.
Thermal melting analysis of the family of 8-aryl-dG adducts has provided insight concerning the effects of adduct size and shape on duplex structure. However, with thermal melting analysis alone, one cannot definitively assign or determine the specific scenarios accounting for the resulting stabilization or destabilization imparted by the modification. Further biochemical analysis is required.

3.4.2.2. Circular Dichroism

As previously mentioned, CD spectroscopy is widely used to determine oligonucleotide folding. CD was carried out on all modified duplexes to assess the specific structural effects of each modification. The CD spectra were taken following the thermal melting analysis of each hybrid, in order to confirm duplex formation prior to taking fluorescence measurements.

In all cases the global DNA structure is that of a B-form helix (Figure 3 - 9). This is characterized by the roughly equal positive and negative CD bands at ~275 nm and ~240 nm, respectively, with a crossover at ~260 nm. The lack of an induced circular dichroism band (ICD) for all adducts except PydG renders the CD spectra almost indistinguishable. For related N-linked C8-dG adducts, ICD has been applied as a diagnostic tool for predicting syn versus anti conformation; a positive ICD signal indicates an S- or W-type duplex conformation, with the adduct residing in the chiral environment in the syn conformation. The Manderville group has previously observed a strong positive ICD signal at ~ 320 nm for the Bth dG adduct incorporated in a decanucleotide (10-mer), hybridized to its complementary strand with the nucleotide G opposite the adduct. Conversely, the CD spectrum of the same 10-mer hybridized to its complementary strand with the nucleotide C opposite the adduct did not exhibit an ICD signal, suggesting it did not reside in the chiral environment of the DNA helix. It is important to note that the absence of an ICD signal is not necessarily conclusive evidence for anti versus syn adduct preference, as the sensitivity of CD is low at the single nucleoside level; it is possible that a resolved ICD band caused by individual bases may be obscured by the signal from the remainder of the molecule.
Therefore, other methods of analysis must be applied in order to confirm the glycosidic bond conformation of the 8-aryl-dG adducts.

Figure 3-9: CD Spectra for mNarI duplexes. (a) NarI(X):NarI'(C) (b) NarI(X):NarI'(G) (c) NarI(X):NarI'(10) (d) NarI(X):NarI'(THF). Black dotted = dG; Black solid = $^{13}$dG; Orange = $^{15}$dG; Red = $^{18}$dG; Green = $^{19}$dG; Blue = $^{10}$dG; Purple = $^{16}$dG.
The CD spectra for the NarI(X):NarI(C) display significant decreases in the positive CD band at 275 nm for all C8-aryl-dG adducts compared to the unmodified duplex. This type of spectral change has been observed during the denaturation of DNA,\textsuperscript{91} or under conditions which induce a change in the winding angle of the DNA resulting in a more coiled helix, with less optimized stacking interactions, but with retention of the base-pairing interactions and backbone conformation.\textsuperscript{92,93} Interestingly, the PydG adduct displayed an ICD with positive amplitudes from approximately 310 nm to 500 nm. This is indicative of the PydG adduct maintaining a syn glycosidic orientation in the duplex, arranging the pyrenyl moiety within the base stack. Most likely, this will cause the opposing C base to flip out, which would also account for the significant duplex destabilization apparent from $T_m$ analysis ($\Delta T_m = -20^\circ$C). Recall that PydG resulted in the largest destabilization noted. It is expected that the S-conformer is slightly more destabilizing than the W-conformer, as it requires the opposing base to also obtain a syn-glycosidic orientation. It seems reasonable to suggest that the PydG adduct would prefer an S-conformation over the W-conformation due to its large $\pi$-surface area and associated lipophilicity.

Conversely, the CD spectra of the NarI(X):NarI(G) showed minimal changes between the modified and unmodified duplexes with regard to the amplitudes of the bands at 240 nm and 275 nm. This type of CD behaviour can be interpreted as the retention of the global and local DNA structure at the site of the lesion between modified and unmodified helices. The largest change in amplitudes resulted from the $^0$dG modification, for which the amplitude at 275 nm was lower than the other modifications; $^0$dG, as discussed, provides little opportunity for effective $\pi$-stacking due to its lack of “length” and is destabilizing due to its “horizontal” bulkiness. The CD data correlate with the $T_m$ analysis, showing that $^0$dG is the most destabilizing adduct within the NarI(X):NarI(G) duplex. Once again the PydG moiety displayed an ICD. Interestingly, it is distinct from the signal produced in the duplex with NarI(C), as it fails to return to the baseline at
500 nm (all 
\textsuperscript{Py}dG CD spectra are overlaid in Figure 3 - 10). It seems likely that the 
\textsuperscript{Py}dG adduct again has adopted a \textit{syn} orientation, this time placing the pyrenyl moiety in the minor groove, allowing for Hoogsteen H-bonding to occur with the opposing G base. This correlates with the \( T_m \) values as \( \text{NarI}(\text{Py}dG):\text{NarI}(G) \) is approximately 6°C more stable than the \( \text{NarI}(\text{Py}dG):\text{NarI}(C) \).

![Figure 3 - 10: CD spectral overlay of all \textsuperscript{Py}dG mNarI duplexes, for complementary strands purple = \text{NarI}(\text{THF}), blue = \text{NarI}(10), red = \text{NarI}(C), green = \text{NarI}(G).](image)

For the \( \text{NarI}(X):\text{NarI}(10\text{mer}) \) duplex oligonucleotides, the CD spectra displayed characteristics of normal B-form DNA with roughly equal positive (\( \approx 275 \text{ nm} \)) and negative (\( \approx 240 \text{ nm} \)) bands, and a crossover at \( \approx 260 \text{ nm} \). Amplitudes at 275 nm were variable, with \( \text{Fur}dG \) and \( \text{CNP}dG \) exhibiting similar heights to the unmodified sequence. The amplitude at 275 nm was diminished for \( \text{Bth}dG, \text{Ph}dG, \text{Py}dG \) and \( \text{O}dG \) modifications. It is hard to interpret these results, as the duplex is inherently distorted, by design, containing a bulge. Once again the \( \text{Py}dG \) adduct displayed the only ICD signal, visually distinct from the \( \text{NarI}(C) \) and \( \text{NarI}(G) \) duplexes as the amplitude was much larger. CD could not definitively distinguish the conformational preference.
of the adducts within this 12-mer:10-mer duplex other than showing that the \( \text{Py} \)dG adduct is once again within the base stack.

The CD spectra for \( \text{NarI}(X):\text{NarI}'(\text{THF}) \) duplexes were very similar to the \( \text{NarI}(X):\text{NarI}'(\text{C}) \) duplexes. The ICD band for \( \text{Py} \)dG was very similar to the \( \text{NarI}(\text{C}) \) scenario, suggesting that the assignment of a base displaced S-conformation for \( \text{NarI}(\text{Py} \text{dG}):\text{NarI}(\text{C}) \) is indeed correct. It seems reasonable to suggest that for the \( \text{NarI}(\text{Py} \text{dG}):\text{NarI}'(\text{THF}) \) duplex the highly lipophilic pyrenyl moiety would prefer a location within the base stack (S-conformation) over one in the minor groove (W-conformation), which would expose it to an aqueous environment. To accommodate the pyrene within the stack, the C base is flipped out, which will have an energetic penalty associated with the anti-syn change in orientation. Indeed, the \( T_m \) data support this, as the \( \text{NarI}(\text{Py} \text{dG}):\text{NarI}'(\text{THF}) \) duplex is 2°C more stable than the \( \text{NarI}(\text{Py} \text{dG}):\text{NarI}'(\text{C}) \) duplex. Moreover, due to the \( \pi \)-surface area of the pyrene ring, its placement within the base stack results in more effective stacking interactions, which has resulted in a larger amplitude at 275 nm relative to all of the other \( \text{NarI}(X):\text{NarI}'(\text{THF}) \) duplexes.

**3.4.3. Conformational preferences of adducts from MD simulations**

MD simulations and free energy calculations were carried out by Purshotam Sharma at the University of Lethbridge, Lethbridge, Alberta. These simulations provide insight into the preferences of the \( \text{Fur} \text{dG}, \text{Ph} \text{dG} \) and \( \text{Q} \text{dG} \) adducted nucleosides within duplexes to adopt an anti or syn conformation. The images presented in Figure 3 - 11 and Figure 3 - 12 were provided by Purshotam Sharma for a joint publication between the Manderville and Wetmore laboratories and are presented here to aid in the description of such systems. Computational details were prepared by Purshotam Sharma and can be found in the general experimental section of Appendix A.
Figure 3 - 11: Lowest energy conformations (within 15 kJ mol$^{-1}$ of free energy) for the studied adducts paired against cytosine. The relative free energies of two competing conformations of Fur$_2$dG are provided in bold in parentheses.

Figure 3 - 12: Lowest energy conformations (within 15 kJ mol$^{-1}$ of free energy) for the studied adducts paired against guanine. The relative free energies of two competing conformations of Q$_2$dG are provided in bold in parentheses.
3.4.4. Fluorescence Response of 8-Aryl-dG Probing Adducts

While still fluorescent, all mNarI oligonucleotides exhibited significantly quenched emission intensity compared to the free modified nucleosides. This large decrease in fluorescence intensity upon oligonucleotide incorporation is similarly observed for many 8-aryl-dG analogues and other fluorescent nucleobase analogues, including 2AP.\textsuperscript{68,74} Several mechanisms have been proposed for fluorescence of 2AP, including base-stacking effects,\textsuperscript{68} and photoinduced electron transfer.\textsuperscript{94}

The excitation spectra of the modified NarI oligonucleotides did display features distinct from the spectra of the free modified nucleosides. In all spectra, a shoulder in the excitation band was present between ~ 270 and 280 nm. In particular, this shoulder was very clearly visible in spectra of QdG, \textsuperscript{CNP}dG, \textsuperscript{Bth}dG and \textsuperscript{Py}dG within NarI. This peak occurs at the red edge of DNA absorbance and can be ascribed to charge transfer from the natural DNA nucleobases to the adduct.\textsuperscript{75} Similar bands have been observed in the emission spectra of 2-AP modified oligonucleotides, in the 260 – 270 nm region.\textsuperscript{95}
Figure 3 - 13: Excitation and emission spectra of 8-aryl-dG mNarI oligonucleotide duplexes. (a) $PdG$; (b) $FurdG$; (c) $BthdG$; (d) $PydG$; (e) $CNPhdG$; (f) $QdG$. Red = NarI'(C); Green = NarI'(THF); Blue = NarI'(G); Purple NarI'(10); Black ssDNA.
Table 3 - 2: Tabulated excitation and emission spectral information for 8-aryl-dG mNarI oligonucleotide duplexes.

<table>
<thead>
<tr>
<th>NarI(X)</th>
<th>NarI'(N)</th>
<th>λex (nm)</th>
<th>Δ λex</th>
<th>λem (nm)</th>
<th>Irel (em)</th>
<th>Δ λem</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhdG</td>
<td>G</td>
<td>300</td>
<td>10</td>
<td>391</td>
<td>0.28</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>302</td>
<td>12</td>
<td>394</td>
<td>0.38</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>296</td>
<td>6</td>
<td>395</td>
<td>0.80</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>297</td>
<td>7</td>
<td>398</td>
<td>0.79</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>290</td>
<td>/</td>
<td>395</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>FurdG</td>
<td>G</td>
<td>312</td>
<td>0</td>
<td>382</td>
<td>0.48</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>314</td>
<td>2</td>
<td>381</td>
<td>0.45</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>306</td>
<td>-6</td>
<td>380</td>
<td>2.97</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>305</td>
<td>-7</td>
<td>383</td>
<td>2.21</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>312</td>
<td>/</td>
<td>381</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>BthdG</td>
<td>G</td>
<td>326</td>
<td>1</td>
<td>416</td>
<td>0.67</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>334</td>
<td>9</td>
<td>422</td>
<td>1.45</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>317</td>
<td>-8</td>
<td>415</td>
<td>1.05</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>329</td>
<td>4</td>
<td>415</td>
<td>0.58</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>325</td>
<td>/</td>
<td>416</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>PydG</td>
<td>G</td>
<td>363</td>
<td>12</td>
<td>459</td>
<td>0.61</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>362</td>
<td>11</td>
<td>454</td>
<td>0.85</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>362</td>
<td>11</td>
<td>455</td>
<td>0.61</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>370</td>
<td>19</td>
<td>451</td>
<td>0.97</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>351</td>
<td>/</td>
<td>456</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>QdG</td>
<td>G</td>
<td>321</td>
<td>7</td>
<td>480</td>
<td>1.55</td>
<td>-6</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>318</td>
<td>4</td>
<td>483</td>
<td>0.41</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>325</td>
<td>11</td>
<td>473</td>
<td>4.92</td>
<td>-13</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>326</td>
<td>12</td>
<td>476</td>
<td>2.24</td>
<td>-10</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>314</td>
<td>/</td>
<td>486</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>CNPhdG</td>
<td>G</td>
<td>332</td>
<td>5</td>
<td>464</td>
<td>2.28</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>327</td>
<td>0</td>
<td>457</td>
<td>1.04</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>328</td>
<td>1</td>
<td>457</td>
<td>3.92</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>THF</td>
<td>333</td>
<td>6</td>
<td>453</td>
<td>2.00</td>
<td>-9</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>327</td>
<td>/</td>
<td>462</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>
The excitation and emission spectra obtained for $^{\text{Ph}}\text{dG}$ modified $\text{NarI}$ oligonucleotides are presented in Figure 3 - 13 (a). In terms of relative fluorescence intensity, all duplexes were quenched in comparison to the single strands, a result that can be attributed to the increased stacking the adduct experiences within a duplex environment compared to ssDNA. The $^{\text{Ph}}\text{dG}$ adduct has been shown to exhibit quenched fluorescence emission in non-polar environments (Table 3 - 1) and for increasing solvent viscosity (work done by Dr. Katie Rankin). The overall trend for $I_{\text{ed}}$ increased in the order $\text{NarI}'(\text{G})$ (0.28) $< \text{NarI}'(\text{C})$ (0.38) $< \text{NarI}'(\text{THF})$ (0.79) $\approx \text{NarI}'(10)$ (0.80). For hybridization to the complement $\text{NarI}'(\text{G})$ $T_m$ analysis suggested a syn orientation of the adduct is obtained, for which a W-type conformation and Hoogsteen H-bonding is expected. This is further supported by the MD simulations, which favour the syn orientation by 37 kJ/mol opposite dG; the lowest energy conformation is presented in Figure 3 - 12. Conversely, $^{\text{Ph}}\text{dG}$ is expected to prefer an anti orientation for $\text{NarI}'(\text{C})$, with a B-type duplex structural motif and W-C H-bonding; this has been corroborated by MD simulations favouring the anti orientation by 26 kJ/mol opposite dC, as presented in Figure 3 - 11. These predictions seem to be supported by the fluorescence data, as the $\text{NarI}'(\text{G})$ is blue shifted and quenched to a higher extent than the $\text{NarI}'(\text{C})$. This is indicative of the ary moiety being in a less polar, aqueous environment, which coincides with the hypothesis that the anti-conformer was preferred in the $\text{NarI}'(\text{C})$ and the syn conformer was preferred in the $\text{NarI}'(\text{G})$. $\text{NarI}'(10)$ and $\text{NarI}'(\text{THF})$ were both more emissive than the $\text{NarI}'(\text{C})$ and $\text{NarI}'(\text{G})$ duplexes. This is consistent with a lack of H-bonding for these situations, which could result in a less restrained environment (although base stacking interactions would be expected to increase rigidity); H-bonding interactions have been shown to result in quenched fluorescence for many 8-aryl-dG analogues (Dr. Katie Rankin’s thesis, unpublished). One would expect the adduct to be in a syn orientation for these scenarios, sequestered from the aqueous environment and positioned within the duplex. Emission intensity was nearly identical for these scenarios, consistent with them both obtaining a syn orientation and being in a similar environment. The fact that the $\text{NarI}'(\text{THF})$ duplex exhibited a slightly red
shifted emission compared to the *NarI*(10) suggests that it is more exposed to the aqueous environment, which may be the result of no complementary nucleobase being present across from the modification (i.e. a “hole” in the duplex allows more interaction with the aqueous environment).

The excitation and emission spectra obtained for *Fur*G modified *NarI* oligonucleotides are presented in Figure 3 - 13 (b) and provided some interesting insight into the syn/anti conformational equilibrium within the duplex. Specifically, the *NarI*(C) and *NarI*(G) were both quenched with regard to the ssDNA, indicating that they were engaged in more H-bonding interactions than the ssDNA. Furthermore, *NarI*(C) and *NarI*(G) duplexes both exhibited nearly identical emission spectra in terms of intensity and emission wavelengths. Of all modifications studied, the *Fur*G adduct has been shown to have the lowest barrier to rotation about the glycosidic bond, which gives rise to an equilibrium of conformations. Calculations performed by Purshotam Sharma of the Wetmore laboratory (University of Lethbridge, Lethbridge, Alberta, Canada) have indicated that *Fur*G exhibited the smallest *anti*/syn energy difference of all adducts studied opposite C (14 kJ mol$^{-1}$) and opposite G (24 kJ mol$^{-1}$) which suggests greater accessibility to both conformations for this adduct. The calculated lowest energy conformations for *Fur*G opposite C and G were presented in Figure 3 - 11 and Figure 3 - 12, respectively. Clearly, fluorescence is in agreement with the simulations, as the emission of *Fur*G indicates a very similar scenario for the adducts within *NarI*(C) and *NarI*(G) duplexes, namely, an equilibrium consisting of relatively even syn and anti orientations of *Fur*G within each duplex. *NarI*(10) and *NarI*(THF) were more emissive than the ssDNA. Once again in these situations there is no opportunity for H-bonding of *Fur*G with a complementary base. *NarI*(10) exhibits larger emission intensity due to the adduct not engaging in effective stacking interactions as $T_m$ analysis showed destabilization for the bulge induced by *NarI*(10).
Previous research conducted by the Manderville group showed that the general spectral features of a $^{\text{Btb}}$dG modified 10-mer oligonucleotide remained unchanged upon hybridization to a complementary strand with C opposing the adduct.\textsuperscript{75} As already discussed, in this duplex, the $^{\text{Btb}}$dG adduct was determined to be in the anti conformation, with supporting evidence from thermal melting, CD and MD analysis. In the anti conformation, the C8-heteroaryl moiety would be located in the major groove and exposed to the polar, aqueous environment surrounding the helical structure, while the syn-conformer is located in a non-polar environment in the minor groove. $^{\text{Btb}}$dG provides some photophysical data that can justify conformation within the NarI recognition sequence; excitation and emission spectra were presented in Figure 3 - 13 (c). The NarI\textsuperscript{(C)} strand is the most emissive ($I_{\text{rel}} = 1.45$) and red-shifted ($\Delta \lambda_{\text{em}} = 6$ nm) which is consistent with the adduct attaining an anti conformation and being exposed to the polar, aqueous environment. NarI\textsuperscript{(G)} and NarI\textsuperscript{(THF)} on the other hand show the largest degree of quenched emission intensity; the adduct is in a syn conformation here, which occupies a less polar environment within the duplex and participates in larger amounts of stacking interactions. Interestingly, the NarI\textsuperscript{(G)} is slightly more emissive than the NarI\textsuperscript{(THF)}, which could be a result of the NarI\textsuperscript{(G)} duplex placing the probe in a W-type conformation (more polar environment) as opposed to an S-type conformation for the NarI\textsuperscript{(THF)}. The NarI\textsuperscript{(10)} shows similar intensity to the ssDNA ($I_{\text{rel}} = 1.1$) and is blue-shifted relative to the NarI\textsuperscript{(C)}, indicative of a syn orientation for the $^{\text{Btb}}$dG adduct.

Push-pull C8-aryl-dG derivatives are highly quenched in the presence of H\textsubscript{2}O and purines, and have the ability to exhibit dual fluorescence, a situation that would provide more specific probing potential within duplex DNA environments. \textsuperscript{Q}dG has also been shown to be sensitive to rigidity as fluorescence emission is enhanced with increasing solvent viscosity. Unfortunately, \textsuperscript{Q}dG failed to exhibit dual fluorescence within any duplex studied. Excitation and emission spectra are presented in Figure 3 - 13 (f). The $I_{\text{rel}}$ values for \textsuperscript{Q}dG mNarI oligonucleotides
increase in the order of $NarI'(C)$ (0.41) $< ssDNA$ (1) $< NarI'(G)$ (1.55) $< NarI'(THF)$ (2.24) $< NarI'(10)$ (4.92). The fluorescence data suggest that, when opposite C, the $QdG$ adduct is *anti*, in a B-type conformation. This conformation exposes the quinoline moiety to the aqueous environment, which results in the largest quenching of emission intensity. $QdG$ has been calculated to favour the *anti* orientation opposite C by 25 kJ/mol; this was unexpected, as the calculations suggest some *syn* conformation could be attainable. The enhanced stabilization of the *syn* conformation when $QdG$ is paired opposite cytosine can be correlated with the greater persistence of two Hoogsteen hydrogen bonds (77% and 53% occupancy). This Hoogsteen interaction was not as persistent as in the case for the $FurdG$ adduct (which displayed ~70% occupancy of both H-bonds). However, the $QdG$ was calculated to favour a W-type structure that contained two Hoogsteen hydrogen bonds between N7 and O6 of $QdG$ and the amino group of the opposing C. Due to the large size of the $QdG$ adduct, it seems as though attainment of this W-type structure would still result in significant exposure to the aqueous environment, which would result in similar levels of fluorescence quenching. Opposite G, MD simulations point to the dominant orientation being *syn*, which is favoured by 76 kJ/mol. Here, the quinolyl moiety would be expected to be in a W-conformation and somewhat sequestered from the solvent, which results in larger $I_{rel}$ values than that of $NarI'(C)$. Interestingly, there were two low energy *syn* orientations obtained by $QdG$ via rotation around the C8-aryl bond (Figure 3 - 12). It can be seen that, with its large twist angles and steric bulk, $QdG$ substantially distorts the duplex and itself is still exposed to an aqueous environment. Opposite the abasic site, the $QdG$ is expected to remain in a *syn* orientation, and most likely adopt an S-conformation. From its emission intensity, the quinolyl moiety appears to be less exposed to solvent, as $I_{rel}$ was larger than either $NarI'(C)$ or $NarI'(G)$. Furthermore, the emission of $QdG$ was slightly blue-shifted opposite the abasic site compared to opposite G, suggesting that $QdG$ adopts an S-conformation within the non-polar base stack. Of the duplexes studied, $NarI'(10)$ had the largest emission intensity and was blue-shifted. The $T_m$ data suggests that the $QdG$ could be present within the bulge, as no destabilization for this
sequence was observed. This agrees with the fluorescence emission intensity as the quinolyl moiety would be sequestered from the aqueous environment.

PydG exhibited quenched fluorescence intensities for all duplexes in comparison to the ssDNA. The relative emission intensity increased in the order: Narl(G) (0.61) = Narl(10) (0.61) < Narl(C) (0.85) < Narl(THF) (0.97) < ssDNA. ICD bands were observed for all Narl duplexes, which indicated that the pyrenyl moiety was within the base stack. Narl(G) was hypothesized to be in a W-type conformation. In this conformation the pyrenyl moiety would be in the minor groove, partially exposed to water, which would account for its strongly quenched fluorescence. The fact that Narl(10) exhibited the same Irel as Narl(G) suggests that the pyrenyl moiety is in a similar environment. The Narl(10) duplex is locally distorted at the adducted site, which is expected to allow some exposure to solvent. Narl(C) was hypothesized to be in an S-type conformation, where the pyrenyl moiety is directly within the base stack. The Irel values support this hypothesis as they are larger than the Narl(G) scenario, which with its W-type conformation is expected to be slightly more exposed to water. The observed Irel for Narl(THF) also suggests the S-type conformation, which has the pyrenyl moiety sequestered from water. The fact that all duplexes are less emissive than their ssDNA counterparts suggests that some form of association or sequestering of the pyrenyl moiety occurs in the ssDNA scenario. Pyrene can form excimers or excited dimers, which reduces solvent exposure.96,97 Here, the pyrene moieties interact with one another in a very tight association, in some cases to avoid exposure to solvent. This scenario would be more prevalent in ssDNA, where the flexibility of the oligonucleotide is greatest. Alternatively, ssDNA may fold around the PydG monomer, thus reducing aqueous exposure of its large, hydrophobic pyrene moiety.

CNPhdG is another member of the D-A 8-aryl-dG probes, like PydG and QdG, it will be quenched in the presence of neighbouring purines and H2O. Probe performance within the Narl duplexes resulted in increasing emission intensities as follows: ssDNA (1) = Narl(C) (1.04) <
\( \text{NarI}(\text{THF}) (2.0) < \text{NarI}(\text{G}) (2.28) < \text{NarI}(10) (3.92) \). Opposite C, the \( \text{CNPh} \text{dG} \) adduct is in an \textit{anti} orientation, with a B-type conformation, which exposes it to the aqueous environment and quenches emission intensity. Opposite G, the \( \text{CNPh} \text{dG} \) is in a \textit{syn} orientation, and expected to be in a W-conformation, resulting in larger \( I_{\text{rel}} \) values than the of \( \text{NarI}(\text{C}) \). Opposite the abasic site the \( \text{CNPh} \text{dG} \) is expected to remain in a \textit{syn} orientation and most likely adopt an S-conformation. Emission intensity supports the \( \text{CNPh} \text{dG} \) being in an S-conformation, as \( I_{\text{rel}} \) was larger than \( \text{NarI}(\text{C}) \), and the emission was blue shifted compared to \( \text{NarI}(\text{G}) \). \( \text{NarI}(10) \) had the largest emission intensity of the duplexes studied; this established that the \( \text{CNPh} \text{dG} \) is present within the bulge and sequestered from the aqueous environment.

### 3.5. Conclusions

Aromatic mutagens undergo metabolic activation to produce aryl radical species that react covalently at the C8-site of dG to produce carbon-linked C8-aryl-dG adducts. Examples include PAHs, arylhydrazines, estrogens and phenolic toxins that are precursors of radical cations, phenyl radicals, radical anions and phenoxy radicals. The resulting biaryl C-linked C8-dG adducts differ greatly in C8-aryl ring size and shape (i.e. phenyl versus benzo[a]pyrene). Such biaryl C8-aryl-dG adducts are more rigid than their N-linked counterparts and the nucleosides strongly favor the \textit{syn} conformation. The biaryl adducts also lack planarity and exhibit a high degree of twist between the nucleobase and C8-aryl group to reduce steric interactions. Such factors result in profound differences when comparing the N-linked-dG and C8-aryl-dG adducts within duplex environments.

Arylhydrazines that produce phenyl radical intermediates are mutagenic in the Ames test. The simplest derivative, \( \text{Ph} \text{dG} \), has been incorporated site-specifically into an oligonucleotide substrate. High-fidelity polymerases, including replicative \textit{E.coli} Klenow fragment and polo, preferentially insert correct cytosine and misinsert some adenine opposite \( \text{Ph} \text{dG} \) \textit{in vitro}, but are strongly impeded by the adduct. Peroxidase-mediated oxidation of benzo[a]pyrene produces
radical cations that are mutagenic in yeast. However, the C8-benzo[a]pyrene-dG adduct has not been incorporated into an oligonucleotide substrate. The phenolic toxin ochratoxin A (OTA) is mutagenic in rat kidney, and affords a C-linked C8-dG adduct in rat kidney DNA. However, a correlation between OTA-mediated mutagenicity and C8-dG adduct formation has not been established. Overall, little is known about the biological implications of C-linked C8-dG adduct formation.

Biaryl C8-aryl-dG adducts are also highly fluorescent. Fluorescent nucleobase analogues are important tools in biology, biochemistry and biophysical chemistry. The family of “probing adducts” described have been shown to be useful for defining adduct conformation in duplex DNA, which can play a critical role in mutagenic outcome. The family of C-linked C8-dG probing adducts studied contained varying C8-aryl ring size, shape, and electronics to provide a systematic analysis for understanding the possible biological consequences stemming from C-linked C8-dG adduct formation. The C-linked adducts were incorporated to determine their conformational preference and thermal impact on duplex stability when paired opposite N (N = C, G and THF). This work demonstrates the relationship between C8-aryl ring size and adduct conformation, which has implications regarding targeted and semi targeted mutagenicity. It is evident that the rules established for mutagenic potency of N-linked C8-dG adducts cannot be extended to the C-linked C8-dG counterparts; due to the absence of the flexible tether separating the dG moiety from the bulky C8-aryl group. The interactions of these 8-aryl-dG adducts with the Klenow fragment and Dpo4 polymerases has been performed by Anne Verway and is awaiting publication.
3.6. References


(39) Gupta, R. C.; Dighe, N. R. *Carcinogenesis* 1984, 5, 343.


(49) Lindsley, J. E.; Fuchs, R. P. *Biochemistry* 1994, 33, 764.


(53) Toth, B.; Gannett, P. *Mycopathologia* 1993, 124, 73.


Chapter 4: Harnessing G-tetrad Scaffolds within G-Quadruplex Forming Aptamers for Fluorescence Detection Strategies
Note on Reproduction of Work

Portions of the research presented in this chapter have been previously published as a first-authored journal articles, with the corresponding citations provided below, and are reproduced herein with permission from the respective journal.


4.1. BIG PICTURE

Much work has gone into the chemical modification of aptamers, with the aim of extending their utilization in therapeutic and biotechnological applications. One key challenge to their successful application is transforming the aptamer-binding events into physically detectable signals. To meet the challenge, a number of methods have been developed, most of which involve fluorescence-based detection. It remains desirable to have a more general method to construct sensors for any aptamer of interest. Preferably, the sensors should be simple to design, easy to operate, give a fast response and have a minimal consumption of material.

Extensive efforts to incorporate fluorescent modifications within aptamers have been reported within the literature and have enjoyed success. However, these modifications are almost always incorporated at the terminus or loop regions of the quadruplex structure. This limits applicability, since each modification is specific to the given aptamer structure. We hypothesized that utilization of C8-aryl-dG analogues, which are sensitive to their microenvironment, and have shown to be good probes for detection within duplex environments, would provide an attractive avenue to pursue.

In this chapter the incorporation of various modified 8-aryl-dG nucleobases into the thrombin binding aptamer (TBA) is discussed. Structure-activity relationships are generated and described; in general the modifications are well tolerated by the structure, although this is somewhat sequence-dependent. TBA was chosen as a model because of the information in the literature pertaining to its overall structure and interaction with thrombin. Through these fluorescent modifications of TBA, it will be shown that it is possible to monitor duplex quadruplex exchange reactions. Also it is shown that the aptamers retain their ability to bind to thrombin.
Modifying aptamers with C8-aryl-dG fluorescent nucleosides, at positions within G-tetrads, should be applicable to numerous G-quadruplex forming aptamers. This study aims to provide insight into the impact of incorporating fluorescent modified nucleobases within TBA.

4.2. Background

4.2.1. Generation of an Aptamer

Aptamers are generated using Systematic Evolution of Ligands by Exponential Enrichment (SELEX); this refers to an in vitro selection process to identify oligonucleotide sequence spaces with specific activities. Oligonucleotide sequence spaces are very large, as they contain all possible sequences of a given length separated by point mutations. The SELEX process is depicted in Scheme 4-1 and begins with the synthesis of a very large oligonucleotide library consisting of randomly generated sequences of fixed length, flanked by constant 5' and 3' ends that serve as binding sites for PCR primers. For a randomly generated region of length $n$, the number of possible sequences in the library is $4^n$ ($n$ positions with four possibilities –A,T,C or G– at each position).
SELEX is an iterative process (typically 7 to 15 rounds) for searching libraries of partially random sequences for molecules that can bind to a given target. The evolution of the population is driven by the selection conditions; the stringency (concentration, incubation time, washes, etc.) is increased progressively from round to round in order to select candidates exhibiting the highest possible affinity. DNA directed SELEX involves initially subjecting DNA sequences from an input library to a selective step, in which bound molecules are separated from unbound molecules. In this case, DNA sequences that bind tightly to a column with ligand
immobilized on a matrix will survive the wash step (step 2) and carry forward to step 3. The DNA sequences are then specifically eluted (step 3) and collected. Step 4 involves amplification of the sequences using the polymerase chain reaction (PCR). The overall cycle is then repeated until functional molecules dominate the population, at which point they are cloned and sequenced. Overall, the SELEX process generates a pool of DNA sequences with high ligand affinity. If the SELEX process is utilized to select for RNA aptamers, as opposed to DNA aptamers, a polymerase with reverse transcriptase activity is required to generate the DNA analogues required for PCR amplification. This is due to the PCR not being compatible with RNA, once PCR is complete the oligonucleotides are again transcribed to the corresponding RNA for selection rounds.

The large number of possible oligonucleotide sequences and their molecular diversity allow the isolation of aptamers with affinity for a large variety of molecules. High-affinity aptamers have been selected for many molecules of various sizes including organic dyes, amino acids, antibiotics, peptides, proteins, and vitamins. The high affinity of aptamers for their targets is given by their capability of folding upon binding their target molecule: they can incorporate small molecules into their nucleic acid structure or integrate into the structure of larger molecules such as proteins.

Aptamers, which show high affinity of binding and high specificity of target recognition, are sometimes described as “chemical antibodies.” They actually display a number of advantages over antibodies as depicted in Table 4–1.
Table 4 - 1: Comparison between aptamers and antibodies

<table>
<thead>
<tr>
<th>Aptamers</th>
<th>Antibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binding affinity</strong> can be as low as nM or pM</td>
<td><strong>Binding affinity</strong> typically as low nanomolar to picomolar range</td>
</tr>
<tr>
<td>Entire selection is a chemical process carried out in vitro and can therefore target many molecules.</td>
<td>Selection requires an immune response, therefore difficult to raise antibodies non-immunogenic targets.</td>
</tr>
<tr>
<td>Can select for ligands under a variety of conditions for in vitro diagnostics</td>
<td>Limited to physiologic conditions for optimizing antibodies for diagnostics</td>
</tr>
<tr>
<td>Iterative rounds against known target limits screening processes</td>
<td>Screening monoclonal antibodies time consuming and expensive.</td>
</tr>
<tr>
<td>Uniform activity regardless of batch synthesis</td>
<td>Activities of antibodies vary from batch to batch.</td>
</tr>
<tr>
<td>Pharmacokinetic parameters can be changed on demand</td>
<td>Difficult to modify pharmacokinetic parameters</td>
</tr>
<tr>
<td>Wide variety of chemical modifications to molecule for diverse functions</td>
<td>Limited modifications of molecule</td>
</tr>
<tr>
<td>Return to original conformation after temperature insult</td>
<td>Temperature sensitive and undergo irreversible denaturation</td>
</tr>
<tr>
<td>Unlimited shelf-life</td>
<td>Limited shelf life</td>
</tr>
<tr>
<td>No evidence of immunogenicity</td>
<td>Significant immunogenicity</td>
</tr>
</tbody>
</table>

A fair comparison should also mention, however, the main disadvantage of aptamers: DNA and especially RNA are very sensitive to hydrolytic digestion by nucleases, thus requiring highly specific environments for their applications.² Another major disadvantage of aptamers is their lack of diversity in terms of the monomeric units. Aptamers, which are comprised of nucleobase monomeric units, have only 4 possible “building blocks” whereas antibodies are comprised of amino acid monomeric units and have 20 distinct “building blocks” which is clearly an advantage.

All in all it is not surprising that aptamers have become ideal tools for the development of analytical methods and therapeutics. Several reviews are available describing the multiple applications of aptamers.¹³,¹⁴ Aptamers provide real world solutions for molecular detection strategies. SOMAmers™ (Slow Off-rate Modified Aptamers), generated by Somalogic, been used for discovering previously undetected biomarkers for drug discovery, drug development, and clinical diagnostics. The SOMAscan™ assay measures 1129 protein analytes in only 65 µL.

168
of a variety of biological matrices and can quantify proteins that span over 8 logs in abundance with low limits of detection and excellent reproducibility. The advantages of aptamers applied in this field are obvious: they are applicable to a wide range of target analytes that can be used; their relative ease of synthesis keeps cost to a minimum; and their high affinity guarantees good selectivity and separation effects.\textsuperscript{13,14}

4.2.2. Chemical Modification of Aptamers

Chemically modified aptamers have been developed due to intrinsic limitations of regular RNA and DNA oligomers.\textsuperscript{15} Such modifications aim to accomplish the following objectives: (1) impart chemical diversity to the nucleic acid; (2) improve the pharmacokinetic profile of the aptamer, and (3) equip the aptamer with a biophysical signal that detects ligand.\textsuperscript{13,14,16}

4.2.2.1. Imparting Chemical Diversity

As previously mentioned, nucleic acids are built from four different deoxyribonucleotide building blocks and therefore, compared to antibodies, have a more limited chemical diversity. While aptamers may bind their targets through H-bonding, electrostatic, and \( \pi \)-stacking interactions, they typically lack aliphatic, positively charged and extended additional aromatic functional groups that are not involved in H-bond formation or intra-helical stacking interactions.\textsuperscript{16} Typically, this is addressed by equipping the nucleobases with additional chemical groups. It is imperative that these modifications do not interfere with the classical base-pairing interactions; therefore, the modifications are typically restricted to non-W-C or Hoogsteen H-bonding sites. Most commonly, modifications are introduced at the C5 position of pyrimidines. Such modifications do not affect the H-bonding sites of dC and have been found to be the most tolerated modifications accepted by DNA polymerases, which is required for the generation of the aptamer via SELEX.\textsuperscript{17}
Examples of modified nucleosides used in SELEX are depicted in Figure 4 - 1. In work by Latham et al., 5-pentynyl-dU (Figure 4 - 1(a)) was substituted for dT in the generation of an aptamer for thrombin which displayed similar dissociation constants as the previously selected single-strand DNA (ssDNA) aptamer.\textsuperscript{18} Another example that clearly highlights the importance of this strategy was provided by Vaught et al., who selected aptamers for the tumour necrosis factor receptor superfamily member 9 (TNFRSF9).\textsuperscript{19} Benzyl- and imidazolylethyl-modified deoxyuridine analogues (Figure 4 - 1:(b) and (c)) enabled successful aptamer selection, whereas unmodified natural DNA failed to generate aptamers targeting TNRSF9. Researchers from Somalogic have built an entire company with the mission objective to increase the molecular diversity of the nucleobases. They report that, when selecting against over 1000 different proteins, the overall success rate for aptamer selection rose from below 30% using unmodified nucleotides to more than 80% by incorporating C5-modified nucleotides. Their most successful derivative is the 5-trypto-aminocarbonyl-dU (Figure 4 - 1 (d)), which provides enhanced possibilities to form hydrophobic and π-stacking interactions.\textsuperscript{20}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{image.png}
\caption{Figure 4 - 1: Structures of various modified nucleosides used in SELEX}
\end{figure}
4.2.2.2. Improving the Pharmacokinetic Profile of the Aptamer

Improving the pharmacokinetic profile of an aptamer is required for any aptamer intended for therapeutic regimens. Nature has evolved numerous enzymes for the synthesis, amplification, modification and hydrolysis of nucleic acids. Nucleases are responsible for the hydrolysis and are important for the turnover of endogenous nucleic acids, and also as defence mechanisms against pathogenic organisms and viruses. These enzymes are found in virtually every organism. Human body fluids contain large amounts of nucleases, and chemical modifications are used to stabilize the aptamers against degradation. This is typically addressed by substitution at the 2' position (e.g., 2'-O-methyl, 2'-fluoro),\textsuperscript{21,22} phosphate modification (e.g. phosphorothioate)\textsuperscript{22} or the utilization of locked nucleic acids (LNAs).\textsuperscript{23}

4.2.2.3. Aptasensors

With regard to imparting a measurable signal to aptamers, a number of analytical methods have been developed, including optical,\textsuperscript{24} electrochemical,\textsuperscript{25} or acoustic sensors,\textsuperscript{26} signalling aptamers (using beacons or quantum dots)\textsuperscript{27} and aptamer arrays for high-throughput analysis.\textsuperscript{28}

Detection strategies for aptamers are separated into “label” and “label-free” techniques.\textsuperscript{29} The “label” strategy dominates the field of aptasensor development and typically involves conjugating different tags via a linker to the 5' and 3'-ends of the oligonucleotide.\textsuperscript{29} Fluorescent tags are commonly employed and molecular target detection strategies include utilizing molecular beacons, binary probes for fluorescence energy transfer (FRET) and pyrene excimer probes.\textsuperscript{30} In this design, G-quadruplex folding brings the probes together to signal target detection.\textsuperscript{31} Fluorescent “label-free” strategies for G-quadruplex detection typically employ fluorescent molecules that bind specifically to G-quadruplex structures with enhanced emission. Examples include N-methyl mesoporphyrin IX (NMM), zinc phthalocyanine (Zn-DIGP) and protoporphyrin IX (PPIX).\textsuperscript{29}
Fluorescent nucleobase analogs are employed as a “label” technique for aptasensor development. Compared to bulky fluorescent tags and external fluorescent molecules, internal fluorescent nucleobase probes are advantageous because they minimally perturb the nucleic acid structure and can report subtle structural changes with single nucleotide resolution.\textsuperscript{32} Thus, signal amplification can result from changes in the local environment of the probe, as it interacts directly with the molecular target. Alternatively, the probe could respond to changes in aptamer global tertiary structure upon molecular target binding. In this regard, the common fluorescent nucleobase analogue 2-AP has been inserted into loop-positions of G-quadruplex aptamers.\textsuperscript{33-35} The emissive adenine analogue exhibits highly quenched fluorescence in duplex DNA, while in the G-quadruplex, the emission intensity of 2AP can equal the intensity of the free base. \textsuperscript{33-35} This suggests that 2-AP could be used in DNA molecular devices that employ duplex–quadruplex exchange strategies for target detection. However, an issue with the use of 2-AP in G-quadruplex-folding aptamers is that placement of the base is restricted to loop-positions, where the adenine analogue destabilizes the quadruplexes.\textsuperscript{35}

Unfortunately, most nucleobase modifications cannot be introduced during the SELEX process, as the modified nucleotides are not substrates for polymerases and cannot be used by the relevant enzymes either for generating the initial library or for amplifying the oligomers selected. These modifications are typically termed “post-SELEX modifications”, which means that regular purine or pyrimidine residues in a sequence obtained through SELEX will be replaced with modified versions.

4.2.3. Thrombin Binding Aptamer

The thrombin binding aptamer (TBA) was one of the first therapeutic aptamers discovered; it was isolated by SELEX against human thrombin. After five rounds of selection, aptamers were identified that exhibited moderate binding affinities for thrombin ($K_d = 25$-$200$ nmol/L).\textsuperscript{36} The majority of aptamers identified shared a common conserved hexamer GGTTGG.
A 15 mer GGTTGGTGTGGTTGG prolonged a clotting time from 25 to 169s in purified fibrinogen and from 25 to 43 s in human plasma.\textsuperscript{36} Thrombin is a key regulatory enzyme in the coagulation cascade. It is a serine protease, produced from prothrombin through the action of factor Xa. Thrombin converts fibrinogen into fibrin, which is the building block of the fibrin matrix of blood clots.

Studies have elucidated the structure of the aptamer as well as its binding to thrombin’s anion binding exosite I.\textsuperscript{37} The structure of TBA was solved by NMR spectroscopy in the 1990s.\textsuperscript{38,39} TBA folds into an intramolecular anti-parallel G-quadruplex, with two G-tetrads containing alternating \textit{syn}- and \textit{anti}-Gs in a chair like conformation (Figure 4 - 2). The G-quadruplex of TBA forms a 1:1 complex with potassium ion (TBA-K\textsuperscript+\textsuperscript{)}) that is highly stabilized compared to the corresponding TBA-Na\textsuperscript+\textsuperscript{c} structure. TBA and modified variants of TBA have been extensively studied in terms of structure (using X-ray crystallography, NMR and CD), stability and \alpha-thrombin affinity.\textsuperscript{23,40-43} Changes in biophysical properties and anticoagulant activity resulting from modifications of the original TBA sequence and/or changes of cation composition of the buffer have been attributed to different folding topologies and strand stoichiometry.
Figure 4 - 2: (a) Structure of a G-tetrad with M⁺ cation in central core; (b) TBA with syn-G’s colored red and anti-G’s colored blue; (c) Structure of the ternary Thrombin-TBA-Na⁺ complex; (d) Structure of the ternary Thrombin-TBA-K⁺ complex.

Crystallographic structural characterization of the thrombin-TBA complex formed in the presence of Na⁺ or K⁺ was provided by Krauss et. al. Their results indicated that the different
effects exerted by Na\(^+\) and K\(^+\) on the inhibitory activity of TBA were related to a subtle perturbation of a few key interactions at the protein–aptamer interface. The interactions between TBA and \(\alpha\)-thrombin are both hydrophobic and hydrophilic and involve residues of the TT loops with a further contribution of G5. They also documented the effects of the counter ion (Na\(^+\) vs K\(^+\)) on the properties of the thrombin-TBA complex. The thrombin-TBA-Na\(^+\) structure has greater flexibility than the K\(^+\) counterpart. The authors attributed this to the potassium ion being able to “perfectly fit at the center of the cavity between the two G-tetrads and links together all the eight purine O6 atoms in a distorted anti-prism geometry at the expected coordination distances, thus increasing the rigidity and the stability of the whole structure.” \(^{44}\) The smaller sodium ion is partially disordered and occupies two alternative positions, each one biased to one of the two tetrads which resulted in more flexibility to the aptamer.

4.3. Materials and Methods

4.3.1. Oligonucleotide Synthesis and Purification

To study the impact of 8-aryl-dG adducts on quadruplex structure and stability the adducts Fur\(d\)G, Bth\(d\)G, Q\(d\)G, CNPh\(d\)G and Py\(d\)G were incorporated into the 15mer TBA sequence using solid phase oligonucleotide synthesis. Preparation of the fully protected phosphoramidites was performed as outlined previously in Scheme 2 - 5.

Oligonucleotide synthesis was performed on a 1 \(\mu\)mol scale as outlined in Section 2.3.3 and Appendix A. Reversed phase HPLC purification was performed to isolate pure 15mer oligonucleotide as outlined in Appendix A. Typically the HPLC profiles confirmed that the solid phase DNA synthesis was achieved with high levels of efficiency.

4.3.2. Mass Spectrometry Analysis

Following HPLC purification, the collected peak was collected and lyophilized to dryness and analyzed by electrospray ionization mass spectrometry (ESI-MS) as described in Appendix
A. The ESI spectra showed the expected clusters of multiply charged peaks for the modified oligonucleotides, with results of this analysis summarized in Table 4 - 2. Mass spectra for all C8-aryl-dG modified oligonucleotides are included in Appendix C.

Table 4 - 2: ESI-MS analysis of mTBA oligonucleotides

<table>
<thead>
<tr>
<th>mTBA Oligonucleotide</th>
<th>Product Formula</th>
<th>Calc’d mass</th>
<th>Exptl m/z</th>
<th>Exptl Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fur-dG</td>
<td>C_{154}H_{189}N_{57}O_{58}P_{14}</td>
<td>4789.8</td>
<td>[M-7H]^− = 683.3</td>
<td>4790.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]^− = 597.7</td>
<td>4789.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 531.3</td>
<td>4790</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 478.0</td>
<td>4790</td>
</tr>
<tr>
<td>Bth-dG</td>
<td>C_{158}H_{191}N_{57}O_{34}P_{14}S</td>
<td>4855.8</td>
<td>[M-8H]^− = 606.1</td>
<td>4855.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 538.6</td>
<td>4855.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 484.6</td>
<td>4855</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-11H]^{11−} = 440.5</td>
<td>4855.5</td>
</tr>
<tr>
<td>CNPh-dG</td>
<td>C_{157}H_{190}N_{58}O_{58}P_{14}</td>
<td>4824.8</td>
<td>[M-7H]^− = 688.7</td>
<td>4824.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]^− = 602.1</td>
<td>4824.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 535.5</td>
<td>4824.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 481.5</td>
<td>4825</td>
</tr>
<tr>
<td>Q-dG</td>
<td>C_{159}H_{192}N_{58}O_{60}P_{14}</td>
<td>4850.8</td>
<td>[M-8H]^− = 605.5</td>
<td>4851</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 538.1</td>
<td>4850.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 484.2</td>
<td>4851</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-11H]^{11−} = 440.1</td>
<td>4851.1</td>
</tr>
<tr>
<td>Py-dG</td>
<td>C_{166}H_{195}N_{57}O_{60}P_{14}</td>
<td>4923.9</td>
<td>[M-7H]^− = 702.6</td>
<td>4924.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]^− = 614.6</td>
<td>4923.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 546.2</td>
<td>4923.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 491.5</td>
<td>4924</td>
</tr>
<tr>
<td>CNPh, Fur-dG</td>
<td>C_{161}H_{192}N_{58}O_{60}P_{14}</td>
<td>4890.8</td>
<td>[M - 6H]^6− = 814.1</td>
<td>4890.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 7 ]^{−} = 697.8</td>
<td>4891.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 8 ]^{−} = 610.4</td>
<td>4891.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M − 9 ]^{−} = 542.5</td>
<td>4891.5</td>
</tr>
<tr>
<td>Py, Fur-dG</td>
<td>C_{170}H_{197}N_{57}O_{60}P_{14}</td>
<td>4989.9</td>
<td>[M-7H]^− = 712.1</td>
<td>4990.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-8H]^− = 622.9</td>
<td>4990.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^− = 553.6</td>
<td>4990.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10−} = 478.0</td>
<td>4789</td>
</tr>
</tbody>
</table>
Upon confirmation via mass spectrometry solutions were quantified according to procedures outlined in Appendix A. Oligonucleotide solutions were then prepared in one of two buffer systems: the first buffer was 50 mM sodium phosphate buffer pH 7.0 with 0.1M NaCl; the second buffer was 50 mM potassium phosphate buffer pH 7.0 with 0.1M KCl. Each solution was adjusted to a final concentration of 6 μM mTBA with the corresponding 50 mM phosphate buffer pH 7.0, 0.1M MCl (M = Na⁺ or K⁺). The corresponding duplex was also prepared, utilizing the 15mer oligonucleotide complementary sequence 5′-CCAACCACACCAACC-3′ to generate 6 μM duplex concentration (i.e. 6 μM of each strand) in the sodium phosphate based buffer as outlined in Appendix A.

4.4. Results and Discussion

4.4.1. Impact on Structure and Stability

The structure of TBA and the modifications introduced into it are illustrated in Figure 4-3. All modifications are designed as internal probes that should participate in and report on quadruplex formation, which is a necessary step in the binding of TBA to its target. The adducts utilized in this study can be categorized according to their structure (as either single or fused ring systems) or their electronics (no push pull vs. push pull character). An advantage of using modified nucleotides is that, because they are an integral part of the functional quadruplex, little decrease in binding affinity is expected; also, they have high resistance to photobleaching.45
As previously discussed, attachment of an aryl ring to the C8 site of dG shifts the conformational equilibrium of the glycosidic bond from anti to syn, perturbing the W-C base pairing with its pyrimidine partner. This scenario is commonly considered a disadvantage within duplexes, since the modified base cannot participate in W-C H-bonding. However, as will be shown below, this can be beneficial in the TBA quadruplex, where modifications of the syn, Hoogsteen H-bonding Gs are possible. Impact of C8-aryl-dG adduct size and electronics when present in different positions, namely a syn-G within the G-quartet, an anti-G within the G-quartet and a G in the TGT loop (Figure 4 - 3 (left)) will be determined. Previous modifications of this sequence have included 4-thio-2'-deoxyuridine\textsuperscript{46}, LNA (locked nucleic acid)\textsuperscript{23}, 2'-deoxy-isoguanosine\textsuperscript{47}, RNA\textsuperscript{48}, partial inversion of TBA polarity (5'-3' to 3'-5') with an 5'-5' internucleoside linkage\textsuperscript{42} or changes of loop size and sequence.\textsuperscript{49} Unfortunately, most of these
attempts have resulted in destabilization of the TBA quadruplex structures. However, modifications involving 8-Br-dG\textsuperscript{50}, unlocked nucleic acids\textsuperscript{51}, and 2-FANA derivatives\textsuperscript{52} have resulted in thermally stabilized quadruplexes compared to the unmodified sequence.

### 4.4.1.1. Thermal Melting Analysis

$T_m$ experiments were conducted in order to assess the impact of 8-aryl-modifications on the overall stability of the mTBA. Results are tabulated in Table 4 - 3 and depicted in Figure 4 - 4.

<table>
<thead>
<tr>
<th>Nucleobase</th>
<th>Position</th>
<th>Na\textsuperscript{+} Duplex</th>
<th>$\Delta T_m$ Na\textsuperscript{+}</th>
<th>Na\textsuperscript{+} Quadruplex</th>
<th>K\textsuperscript{+} Quadruplex</th>
<th>$\Delta T_m$ K\textsuperscript{+}</th>
</tr>
</thead>
<tbody>
<tr>
<td>dG</td>
<td>/</td>
<td>64.5 ± 0.7</td>
<td>/</td>
<td>24.0</td>
<td>53.3 ± 1.8</td>
<td>/</td>
</tr>
<tr>
<td>Fur dG</td>
<td>5</td>
<td>57.9 ± 1.6</td>
<td>-6.6</td>
<td>40.2 ± 0.5</td>
<td>62.4 ± 0.9</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>59.6 ± 0.9</td>
<td>-4.9</td>
<td>NO $T_m$</td>
<td>42.3 ± 2.1</td>
<td>-11</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>59.7 ± 1.5</td>
<td>-4.8</td>
<td>NO $T_m$</td>
<td>43.8 ± 2.3</td>
<td>-9.5</td>
</tr>
<tr>
<td>CNPh dG</td>
<td>5</td>
<td>54.8 ± 0.5</td>
<td>-9.7</td>
<td>NO $T_m$</td>
<td>60.4 ± 0.5</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>55.4 ± 1.4</td>
<td>-9.1</td>
<td>NO $T_m$</td>
<td>NO $T_m$</td>
<td>/</td>
</tr>
<tr>
<td>Q dG</td>
<td>8</td>
<td>54 ± 1.5</td>
<td>-10.5</td>
<td>NO $T_m$</td>
<td>47 ± 1.9</td>
<td>-6.3</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>50.8 ± 1.6</td>
<td>-13.7</td>
<td>NO $T_m$</td>
<td>NO $T_m$</td>
<td>/</td>
</tr>
<tr>
<td>Bth dG</td>
<td>8</td>
<td>48.2 ± 1.1</td>
<td>-16.3</td>
<td>NO $T_m$</td>
<td>46.9 ± 2.1</td>
<td>-6.4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>55 ± 1.8</td>
<td>-9.5</td>
<td>NO $T_m$</td>
<td>55.2 ± 1.4</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>55.5 ± 1.4</td>
<td>-9</td>
<td>NO $T_m$</td>
<td>30</td>
<td>-23.3</td>
</tr>
<tr>
<td>Py dG</td>
<td>8</td>
<td>54.6 ± 1.6</td>
<td>-9.9</td>
<td>NO $T_m$</td>
<td>45.2 ± 1.7</td>
<td>-8.1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>52 ± 1.7</td>
<td>-12.5</td>
<td>NO $T_m$</td>
<td>38.0 ± 2.4</td>
<td>-15.3</td>
</tr>
<tr>
<td>Fur dG@5; CNPh dG@8</td>
<td>8</td>
<td>49.0 ± 1.5</td>
<td>-15.5</td>
<td>NO $T_m$</td>
<td>31.2 ± 1.0</td>
<td>-22.1</td>
</tr>
<tr>
<td>Fur dG@10; CNPh dG@8</td>
<td>5.8</td>
<td>49.4 ± 0.8</td>
<td>-15.1</td>
<td>20</td>
<td>60 ± 0.2</td>
<td>6.7</td>
</tr>
<tr>
<td>Fur dG@1; Py dG@8</td>
<td>1.8</td>
<td>48.5 ± 2.3</td>
<td>-16</td>
<td>NO $T_m$</td>
<td>45.6 ± 1.1</td>
<td>-7.7</td>
</tr>
<tr>
<td>Fur dG@5; Py dG@8</td>
<td>5.8</td>
<td>42.5 ± 2.1</td>
<td>-22</td>
<td>NO $T_m$</td>
<td>43 ± 1.3</td>
<td>-10.3</td>
</tr>
<tr>
<td>Adduct</td>
<td>Position 5</td>
<td>Position 6</td>
<td>Position 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fur-dG</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BTh-dG</td>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNPh-dG</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-dG</td>
<td><img src="image10" alt="Graph" /></td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Py-dG</td>
<td><img src="image13" alt="Graph" /></td>
<td><img src="image14" alt="Graph" /></td>
<td><img src="image15" alt="Graph" /></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-4: $T_m$ profiles for 8-aryl-dG mTBA oligonucleotides. Solid blue line represents duplex $T_m$; solid red line represents K$^+$ stabilized quadruplex $T_m$; dotted red line represents Na$^+$ stabilized quadruplex $T_m$. 
All modified duplexes were destabilized upon incorporation of a C8-aryl-dG adduct. Averaged over all positions in the sequence, $T_m$ values were lowered by $-5.4^\circ C$ for Fur-dG, $-9.5^\circ C$ for CNPh-dG, $-10.1^\circ C$ for Bth-dG, $-14^\circ C$ for Py-dG and $-14.6^\circ C$ for Q-dG. The extent of destabilization was clearly correlated to the size of the adduct, as the destabilization was greatest for the fused ring systems, opposed to the single ring ones ($Py-dG \approx Q-dG > Bth-dG > CNPh-dG > Fur-dG$). This is in agreement with the studies on mNarI oligonucleotides discussed above, and with published studies on C8-modified-dG adducts, showing that C8-aryl-dG ($\Delta T_m = -6$ to $-17^\circ C$)\textsuperscript{53-55}, C8-pyrenyl-dG ($\Delta T_m = -10^\circ C$)\textsuperscript{56}, and N-linked C8-arylamine-dG ($\Delta T_m = -8$ to $-13^\circ C$)\textsuperscript{57} lesions significantly destabilize duplexes when paired with dC.

The mechanism of destabilization varies slightly among the modifications, but for C linked C8-aryl-dG modifications it is believed that the adducts’ preference for a syn-conformation is the main cause. If the adducts are able to maintain an anti conformation to engage in W-C H-bonding with the complementary C, the destabilizing effects would result from (1) the lipophilic C8-aryl moiety being exposed to an aqueous environment and (2) steric clash generated between the major groove located C8-aryl substituent and the deoxyribose sugar/phosphate backbone. These data agree with DFT calculation predictions for the adducts preference to adopt a syn glycosidic orientation over an anti, as all optimized syn structures are $\sim 25-30$ kJ/mol more stable than their anti counterparts, with the single rings having a smaller energy gap (and hence less destabilization) than the fused ring systems. For 8-aryl-dG adducts in a syn orientation opposite C (as noted for the Py-dG derivative in the mNarI oligonucleotide), the resulting helical distortion and loss of W-C H-bonding would account for the decreased duplex stability and decrease in $T_m$ compared to the unmodified duplex.\textsuperscript{57}

The modified quadruplex data shows both stabilization and destabilization depending on the modification and location within the TBA. Position 5 of the oligonucleotide is a syn-G (Figure 4 - 3) and was therefore expected to be stabilized. This is indeed the case for all
modifications except the bulkiest QdG ($\Delta T_m = -1.7^\circ C$) and PydG ($\Delta T_m = -15.3^\circ C$) adducts. FurdG showed a very significant stabilization of ($\Delta T_m = +9.1^\circ C$) followed by CNPhdG ($\Delta T_m = +7.1^\circ C$) and BhdG ($\Delta T_m = +1.9^\circ C$), respectively. Once again the size of the modification is playing a critical role; since all adducts prefer a syn orientation (and the larger adducts prefer it more than the smaller ones), it seems that steric clash of the C8-aryl moiety with neighbouring bases and the penalty of solvating the lipophilic moiety result in destabilization for the larger ring systems. Interestingly, the only modification that was able to fold into a corresponding quadruplex in Na$^+$ was FurdG@5 which resulted in a significant stabilization of the quadruplex ($\Delta T_m = +16.2^\circ C$).

Modification at position 6 of the quadruplex resulted in destabilization across the board, so much so that the quadruplexes failed to form for all adducts except the FurdG. This is thought to be the result of the large energy penalty required for the adducts to flip from syn to anti, as required for the dG at position 6 in the TBA. Since FurdG is the smallest and least sterically demanding adduct, it was able to attain an anti conformation, but still caused a significant destabilization of the quadruplex ($\Delta T_m = -11^\circ C$).

Position 8 is not within a tetrad, as it is in the TGT loop of the TBA. All modifications at position 8 were tolerated by the TBA structure, yet all modifications were destabilizing. This suggested that this G preferred the anti conformation in the native TBA structure. The overall destabilization caused by the modifications (PydG $\Delta T_m = -22^\circ C$, FurdG $\Delta T_m = -9.5^\circ C$, BhdG $\Delta T_m = -8.1^\circ C$, QdG $\Delta T_m = -6.4^\circ C$ and CNPhdG $\Delta T_m = -6.3^\circ C$) cannot be rationalized based solely on the size of the modification. Here, the adducts’ ability to stack with the neighbouring tetrad seems to play a role, such that stacking compensates for the anti to syn change in glycosidic bond orientation. The push pull derivatives CNPhdG and QdG are less destabilizing than the non push-pull derivatives. It was surprising that the PydG modification was so strongly destabilizing. Possibly, the aryl ring is too large to stack with the terminal tetrad to offset the energetic penalty of attaining a syn-orientation.
We also incorporated multiple modifications within the TBA sequence. The $T_m$ analysis was presented in Table 4 - 3, while $T_m$ profiles are depicted in Figure 4 - 5. In these experiments, the adducts, and their positions within TBA, were selected according to the information generated on the single adduct incorporations. We took advantage of the pronounced quadruplex stabilization observed with $\text{Fur}dG$ at position 5 and therefore was utilized to impart added stability to the quadruplex. $\text{CNPh}dG$, while destabilizing ($\Delta T_m = -6.3^\circ C$), had interesting fluorescent properties at position 8 as outlined in Table 4 - 3. Generating the doubly modified, $\text{Fur}dG/\text{CNPh}dG$ oligonucleotide combined the stability imparted by $\text{Fur}dG@5$ with the interesting fluorescent signals generated by $\text{CNPh}dG@8$; these oligonucleotides proved interesting to study. Because of its large destabilizing effect ($\Delta T_m = -22^\circ C$), $\text{Py}dG$ was used in only a few experiments; it was again combined with $\text{Fur}dG$ for the sake of oligonucleotide stability in order to extend its applicability.
Figure 4 - 5: $T_m$ profiles of various doubly modified mTBA oligonucleotides. (a) CNPh\dG@8/Fur\dG@5; (b) CNh\dG@8/Fur\dG@10; (c) Fur\dG@1/Py\dG@8; (d) Fur\dG@5/Py\dG@8. Solid red line corresponds to TBA-K$^+$ quadruplex $T_m$; Dotted line is TBA-Na$^+$ quadruplex $T_m$; Solid Blue line corresponds to TBA-Na$^+$ duplex.

As expected, modification of an additional position within TBA resulted in further duplex destabilization (Table 4 - 3). Interestingly, the $T_m$ of \Fur\dG@5/CNPh\dG@8 mTBA oligonucleotide was lowered less than expected from the effects of the two individual modifications (i.e. summing the $T_m$ contribution of the individual modifications resulted in a hypothesized $T_m$ of 47.4°C whereas the observed $T_m$ was 49.4°C). This was also the case in the \Fur\dG/Py\dG modified oligonucleotides. Site-specific destabilizations resulted in \Fur\dG@5/Py\dG@8 generating an observed $T_m$ of 42.5°C which is in close agreement with the predicted $T_m$ (addition of individual destabilizations) while \Fur\dG@1,Py\dG@8 had a $T_m$ of 48.5°C which is 6°C more stable than
predicted values. These discrepancies in melting temperature were important to note, as it suggested that larger amounts of destabilization result when DNA contains bulky modifications within a localized portion of the duplex. The duplex structure seems to better cope with pair wise lesions when these are placed distantly rather than in close proximity.

The impact of incorporating two modifications on quadruplex structure was also determined. The $^{\text{Fur-dG/CNPh-dG}}$ oligonucleotides performed as intended. The $^{\text{Fur-dG}}$ at a syn position was expected to stabilize the quadruplex (single incorporation had a $\Delta T_m = +9^\circ C$), while the $^{\text{CNPh-dG}}$ at position 8 was expected to destabilize the quadruplex, since the single incorporation had a $\Delta T_m = -6^\circ C$. The observed $T_m$ of approximately 60$^\circ C$ represented a 6$^\circ C$ increase in stability and was observed with $^{\text{Fur-dG}}$ in any of the positions 1, 5 and 10. $^{\text{Py-dG@8}}$ afforded a $T_m$ of 31.2$^\circ C$ for the single modification; adding $^{\text{Fur-dG@5}}$ resulted in a $T_m$ of 45.6$^\circ C$, which was a much larger stabilization than expected (i.e. adding the $^{\text{Fur-dG@5}}$ contributed a $\Delta T_m = +14.4^\circ C$). Adding $^{\text{Fur-dG@1}}$ also afforded significant stabilization ($\Delta T_m = +11.8^\circ C$), as its $T_m$ was 43$^\circ C$. The larger degrees of stabilization imposed by $^{\text{Fur-dG}}$ when the $^{\text{Py-dG}}$ was present at position 8 may be explained by entropy effects. In an unmodified oligonucleotide, quadruplex formation requires certain nucleosides (position 5 included) to switch their glycosidic orientation from anti to syn. Placing $^{\text{Fur-dG}}$ to position 5 promotes transition to syn and may aid in the overall transformation from ssDNA to quadruplex.

4.4.1.2. Circular Dichroism

As previously mentioned, CD is commonly used to determine oligonucleotide folding. CD was carried out on all modified duplexes and to confirm that the mTBA sequences retain the ability to fold into anti-parallel quadruplexes. Spectra were monitored at 6 $\mu$M concentrations and data is presented in Figure 4 - 6 (duplexes) and Figure 4 - 7 (quadruplexes).
Figure 4 - 6: CD spectra of mTBA duplexes. (a) modification at position 5; (b) modification at position 6; (c) modification at position 8. Black = unmodified TBA; Orange = $^{\text{Py}}$dG mTBA; Red = $^{\text{Fur}}$dG mTBA, Blue = $^{\text{CNPh}}$dG mTBA; Purple = $^{\text{Q}}$dG mTBA; Green = $^{\text{Bth}}$dG mTBA.

All recorded CD spectra for duplexes displayed the characteristic positive long wavelength band or bands between 260–280 nm and a negative band around 245 nm. Interestingly, there was a blue shift of the observable signals for all modified oligonucleotides with the $^{\text{Fur}}$dG, $^{\text{CNPh}}$dG, $^{\text{Bth}}$dG, $^{\text{Q}}$dG and $^{\text{Py}}$dG displaying a prominent peak at approximately 266 nm and a shoulder at 283 nm which overlapped with the unmodified TBA duplex.

Another difference between modified and unmodified duplex spectra can be seen for the $^{\text{Py}}$dG@5 mTBA oligonucleotide (Figure 4 - 6 (a) orange trace). $^{\text{Py}}$dG displayed large increases in amplitude for the characteristic B-DNA conformation and an induced CD band with negative
amplitude at approximately 300 nm and positive amplitude from approximately 325 to 400 nm. Generally, increases in amplitude have been attributed to more effective stacking interactions. The presence of the induced CD band indicates that the pyrene moiety is stacked within the interior of the duplex. This was previously noted for PydG incorporation within the mVarI sequence. Intercalation of PydG within the duplex can result in greater amplitudes in the CD spectra, since the π-surface of the pyrene system is larger than that of dG and increases the stacking interactions of the duplex overall. The absence of an amplitude or induced CD band for the PydG at position 8 suggests these findings are sequence-specific and modulated by the base sequence context. Similar observations have been reported for PydG in the literature by Valis et al. In their experimental design, the PydG was studied in duplexes where the modification was placed between different flanking bases; the sequence was 5′-GCAGTCTN(PyG)NTCACTGA-3′ where N was either A,T,G or C. When PydG was placed between two purines, the CD signal for the corresponding oligonucleotides contained a large negative peak at approximately 305 nm and a broad positive peak from approximately 330 to 410 nm in addition to the characteristic B-DNA signals at 278 nm and 245 nm. In comparison, the CD spectrum of PydG flanked by two pyrimidines had a very small negative peak at approximately 300 nm, which did not contain a positive peak in the 330-400 nm region. Unfortunately, the authors did not interpret or expand on their findings and simply presented their spectra in the supporting information.

It is clear that the induced CD bands for the TBA PydG at position 5 are not as large as the ones depicted by Valis et al. This could be due to the fact that our PydG system is flanked by one T and one G at position 5. It is well known that purines are better able to stack with purines than pyrimidines due to their larger π surfaces. Therefore, a stacked conformation would be the most stable and most prominent when it is surrounded by two purines, less stable when it is surrounded by a single purine and least stable when flanked by two pyrimidines. It is important to note that stacking of the pyrene moiety is dependent on other factors. Although its π-surface is
larger than that of dG, placement of the pyrene moiety within the duplex disrupts the W-C H-bonding interaction that dG shares with its opposing dC base and can be thought of as an energetic penalty; our $T_m$ results corroborate this, as $^{Py}dG$ duplex has a $\Delta T_m = -12.5^\circ C$. The ability of pyrene to stack depends on whether or not the gain in energy from stacking can compete with such a penalty which, in this case, occurs when flanked by purines.

CD spectra for all mTBA oligonucleotides are presented in Figure 4 - 7. As previously mentioned, anti-parallel quadruplexes are characterized by a negative CD band at 260 nm and positive CD bands around 210 nm, 240 nm and 290 nm respectively.
All quadruplexes in the presence of Na\(^+\) ions displayed lower CD intensities than their K\(^+\) counterparts; this can be rationalized in two ways. (1) All Na\(^+\)-stabilized quadruplexes display a
lower $T_m$ than their K$^+$ counterpart; therefore, folded quadruplexes are less abundant. (2) According to crystallographic studies, the Na$^+$-stabilized quadruplex of unmodified TBA is less rigid, which should result in loosened stacking interactions.\textsuperscript{44}

CD spectral interpretation shows that modifications at G5 of TBA is well tolerated, as all signals overlap with the unmodified quadruplex spectra. The peak height at 295 nm relates to the overall size of the modification as $^\text{Py}dG < ^\text{Q}dG < ^\text{Bth}dG \approx ^\text{CNPh}dG < ^\text{Fur}dG$. It is important to note that the 8-aryl moiety is unable to stack with any portion of the quadruplex structure, as it is projected outward into an aqueous environment. It seems reasonable to suggest that larger aryl systems will clash with the oligonucleotide backbone to a higher extent, and result in local distortions of the G-quadruplex, thus affecting the extent of G-tetrad stacking. However, the size of the aryl ring is not the only factor that affects G-tetrad stacking, as evidenced for the $^\text{CNPh}dG$ and $^\text{Bth}dG$ modifications. These modifications display similar amplitudes in their CD spectra even though one would expect the larger $^\text{Bth}dG$ to be more disruptive (as was observed with the $T_m$ analysis). Adduct twist angles $\theta$ may account for this. Below in Figure 4 - 8 is the unmodified TBA quadruplex in the presence of K$^+$ with G5 colored in Red and C8 of G5 highlighted in green. One can see that a larger twist angle $\theta$ between an aryl ring attached at the C8 carbon would result in more steric clash between the two tetrads. This should result in local distortion of the quadruplex, reducing the overall co-planarity (and stacking) of the tetrads. This unfavourable steric clash is minimized as $\theta$ approaches 0, since the adduct is able to attain planarity, and the stacking interactions between tetrads are again maximized.
Fur-dG and CNP-dG modifications were tolerated when placed at the syn-G6 position of TBA in K⁺. This was not the case for Q-dG or Bth-dG; Py-dG was not incorporated at G6, as it was not expected to fold into a quadruplex. Accommodation of these modified bases correlated with their ability to adopt an anti conformation, which prevails in the native sequence at position 6. Molecular dynamic simulations have provided differences in energy for syn-anti conformations for all adducts and show that the Fur-dG and CNP-dG have much smaller energy gaps for obtaining anti structures than the Q-dG Bth-dG or Py-dG. Only the Fur-dG modification seemed to be tolerable in sodium, although no Tₘ’s were afforded for any of these structures.

Modification of G8 of TBA provided insight into how a modification is tolerated in the loop of TBA, i.e. a non tetrad G-modification. Overall, the anti-parallel quadruplex topologies were conserved in potassium and sodium. Depicted in Figure 4 - 9 is the unmodified TBA-K⁺-quadruplex with G8 highlighted in red and C8 in green. It is seen that G8 is normally in an anti conformation. However, unlike the dG positions within the tetrad, G8 will not disrupt quadruplex
folding when it deviates from this conformation. Analysis of the CD spectra shows that modification of G8 with $^{Py}$dG resulted in a reduction of the CD signal amplitude; therefore, $^{Py}$dG was either unable to attain an anti conformation, or else, due to its size, failed to engage in effective stacking with the above tetrad. All other modifications resulted in increased amplitude compared with the unmodified sequence, which may be due to the aryl moiety extending the π-surface of the nucleoside, which should allow for more effective overlap and stacking. The smaller ring systems generated the largest amplitudes in CD signals, as they were more planar and more willing to adopt an anti conformation (i.e. $^{Fur}$dG > $^{CNPh}$dG ~ $^{Bth}$dG > $^{Q}$dG >> $^{Py}$dG).

Figure 4 - 9: Crystal structure of K+-TBA with G8 colored red and C8 of G8 highlighted in green. Adapted from previously reported crystal structure (pdb code 4DII).44

In general, the interpretation of CD spectra has provided a general knowledge on how well the modifications are tolerated within different positions of a quadruplex. This information becomes very useful when trying to design modified aptamers. It can generally be said that 8-aryl-dG modifications should be restricted to those dG positions within the tetrads that prefer syn
orientations; in some cases, with small enough probes, modifications to the dGs within the loop structures are well tolerated. If modifications within the loop structures are required, these findings demonstrate that complementing them with a modification of a separate syn-preferring G will help to stabilize the overall structure.

4.4.2. Probe Performance

Johnson et al. have recently published on the fluorescent properties of 2-AP in loop positions of TBA.\textsuperscript{35} Thymidine → 2AP substitution in TBA resulted in destabilization ($\Delta T_m = -8^\circ C$) of the quadruplex at positions 3, 7, 9 and 12, while substitution at position 4 and 13 resulted in slight stabilization ($\Delta T_m = +2^\circ C$). The fluorescence of 2-AP was also sensitive to the site of incorporation. The authors found that 2-AP in positions 4, 9 and 13 quenched to the same degree as within DNA duplexes. In positions 3 and 12, 2-AP exhibited an approximate 30 fold increase in fluorescence intensity, while position 7 had the highest fluorescent signal which was approximately 70 fold higher. The authors attributed the difference in fluorescence intensity to the probable formation of a T4-T13 base pair (when one T was replaced with 2-AP) across the G-quartet, as evident from ad nuclear Overhauser effect between the imino protons of the nucleobases.\textsuperscript{38} This base pair was proposed to be involved in a strong stacking interaction with the G-quartet, resulting in effective quenching of the 2-AP residues at these positions. The authors determined that T3, T12 and T7 were not interacting with the adjacent nucleotides and were accessible to the solvent, which resulted in higher levels of emission of 2-AP.

This study demonstrated the utility of incorporating fluorescent nucleobase analogues within quadruplexes, as the authors were able to probe local environment and correlate their emission response with structural studies of the oligonucleotide. However, the use of 2-AP in this study was clearly limiting. Firstly, 2-AP is an adenosine derivative (i.e. a nucleobase not naturally present in TBA), it is able to form a base pair with thymidine; such a base pair is not present in unmodified TBA and clearly has the potential to change the overall structure and
dynamics of the G-quadruplex. This was evident from the destabilizing effect observed upon incorporation of 2-AP within TBA, except for at positions T4 or T13 in which the 2-AP residue was able to H-bond with its complementary thymidine.

It would clearly be beneficial to have alternative modifications that are less disruptive to the quadruplex structure. $T_m$ and CD experiments have demonstrated that 8-aryl-dG analogues have the potential to serve this purpose. None of these analogues affect base pairing, and the Fur dG analogue even stabilizes the G-quadruplex structure. Modifying positions within the tetrad of a quadruplex allows for the probes to be present in positions that undergo large changes in their base stacking arrangement upon ligand binding (i.e. single strand environment to a stacked tetrad within the quadruplex). Also, since tetrads are formed universally within G-quadruplexes, their modification should be more applicable to aptamer development compared to modifying the loop structures, which are more limited to participate in stacking interactions.

Many of the fluorescent analogues reported by the Manderville have a quantum yields that are sensitive to their immediate surroundings.$^{54,60-62}$ The solvatochromic properties of the C8-aryl−dG nucleoside adducts incorporated within TBA were presented in Table 3 - 1. The photophysical properties of C8-aryl-dG mTBA oligonucleotides were here examined. Emission and excitation spectra of the adduct-containing TBA oligonucleotides in the duplex and quadruplex state are shown in Figure 4 - 10 with photophysical parameters for each adduct depicted in Tables 4-5 to 4-9.
Figure 4 - 10: Excitation and emission spectra for 6 μM solutions of mTBA oligonucleotides. Solid red line represents modification at position 5 in K⁺; dotted red line represents modification at position 5 in Na⁺; blue line represents modification at position 6; black line represents modification at position 8.
Table 4 - 4: Tabulated excitation/emission data for 6 μM solutions of Fur dG mTBA.

<table>
<thead>
<tr>
<th>Position</th>
<th>Strand</th>
<th>$\lambda_{ex}$ (nm)</th>
<th>Intensity</th>
<th>$\lambda_{em}$ (nm)</th>
<th>Intensity</th>
<th>I$_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Na$^+$ Quad</td>
<td>259.1</td>
<td>127.9</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>291.1</td>
<td>141.7</td>
<td>378.7</td>
<td>140.5</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>315.1</td>
<td>134.8</td>
<td>378.9</td>
<td>131.1</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quad</td>
<td>258.0</td>
<td>151.9</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>291.1</td>
<td>165.9</td>
<td>378.8</td>
<td>163.5</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>315.1</td>
<td>147.5</td>
<td>378.9</td>
<td>141.0</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>13.7</td>
<td>390.2</td>
<td>13.4</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>314.0</td>
<td>24.7</td>
<td>390.0</td>
<td>22.8</td>
<td>/</td>
</tr>
<tr>
<td>6</td>
<td>K$^+$ Quad</td>
<td>254.1</td>
<td>624.0</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>289.1</td>
<td>640.0</td>
<td>381.1</td>
<td>630.0</td>
<td>37.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>324.0</td>
<td>483.0</td>
<td>381.1</td>
<td>461.2</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>17.2</td>
<td>381.2</td>
<td>17</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>315.1</td>
<td>37.4</td>
<td>382</td>
<td>36.8</td>
<td>/</td>
</tr>
<tr>
<td>8</td>
<td>K$^+$ Quad</td>
<td>251</td>
<td>252.9</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>297.1</td>
<td>346.1</td>
<td>381.1</td>
<td>332.0</td>
<td>31.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>315.1</td>
<td>356.0</td>
<td>381.1</td>
<td>347.8</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>10.8</td>
<td>384.0</td>
<td>10.6</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>314.0</td>
<td>19.5</td>
<td>384.0</td>
<td>38.3</td>
<td>/</td>
</tr>
</tbody>
</table>
Table 4 - 5: Tabulated excitation/emission data for 6 μM solutions of \textsuperscript{3}HdG mTBA.

<table>
<thead>
<tr>
<th>Position</th>
<th>Strand</th>
<th>$\lambda_{ex}$ (nm)</th>
<th>Intensity</th>
<th>$\lambda_{em}$ (nm)</th>
<th>Intensity</th>
<th>$I_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Na$^+$ Quad</td>
<td>254.1</td>
<td>495.8</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>290.0</td>
<td>480.7</td>
<td>415.0</td>
<td>491.7</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>331.0</td>
<td>418.4</td>
<td>416.1</td>
<td>438.4</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quad</td>
<td>250.9</td>
<td>496.3</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>291.1</td>
<td>490.9</td>
<td>418.9</td>
<td>485.3</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>327.1</td>
<td>483.9</td>
<td>418.9</td>
<td>485.3</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>175.0</td>
<td>418.9</td>
<td>173.3</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>345.1</td>
<td>419.9</td>
<td>418.9</td>
<td>374.6</td>
<td>/</td>
</tr>
<tr>
<td>6</td>
<td>K$^+$ Quad</td>
<td>250.9</td>
<td>480.5</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>287.0</td>
<td>392.8</td>
<td>420.0</td>
<td>358.9</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>338.0</td>
<td>406.2</td>
<td>420.0</td>
<td>358.9</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>125.0</td>
<td>423.0</td>
<td>125.0</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>333.5</td>
<td>287.1</td>
<td>423.0</td>
<td>251.1</td>
<td>/</td>
</tr>
<tr>
<td>8</td>
<td>K$^+$ Quad</td>
<td>250.0</td>
<td>182.8</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>286.0</td>
<td>162.6</td>
<td>416.1</td>
<td>160.1</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>327.1</td>
<td>190.9</td>
<td>416.1</td>
<td>180.1</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>122.0</td>
<td>421.1</td>
<td>120.4</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>343.0</td>
<td>407.9</td>
<td>421.1</td>
<td>430.8</td>
<td>/</td>
</tr>
</tbody>
</table>
Table 4 - 6: Tabulated excitation/emission data for 6 μM solutions of \( \text{^{CNPP}dG} \) mTBA.

<table>
<thead>
<tr>
<th>Position</th>
<th>Strand</th>
<th>( \lambda_{\text{ex}} ) (nm)</th>
<th>Intensity</th>
<th>( \lambda_{\text{em}} ) (nm)</th>
<th>Intensity</th>
<th>( I_{\text{rel}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Na(^+) Quad</td>
<td>243.0</td>
<td>171.5</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>290.0</td>
<td>138.7</td>
<td>460.0</td>
<td>137.6</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>330.0</td>
<td>129.6</td>
<td>460.0</td>
<td>122.9</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>K(^+) Quad</td>
<td>250.0</td>
<td>108.0</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>290.0</td>
<td>86.7</td>
<td>461.0</td>
<td>85.0</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>330.0</td>
<td>79.9</td>
<td>461.0</td>
<td>77.2</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>31.1</td>
<td>460.0</td>
<td>30.5</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>335.0</td>
<td>77.9</td>
<td>460.0</td>
<td>69.1</td>
<td>/</td>
</tr>
<tr>
<td>6</td>
<td>K(^+) Quad</td>
<td>247.0</td>
<td>320.3</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>285.9</td>
<td>217.0</td>
<td>460.0</td>
<td>215.1</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>339.1</td>
<td>216.2</td>
<td>460.0</td>
<td>197.3</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>37.7</td>
<td>467.0</td>
<td>36.9</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>328.0</td>
<td>113.4</td>
<td>467.0</td>
<td>110.5</td>
<td>/</td>
</tr>
<tr>
<td>8</td>
<td>K(^+) Quad</td>
<td>243.0</td>
<td>152.0</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>284.1</td>
<td>129.2</td>
<td>453.0</td>
<td>127.5</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>327.1</td>
<td>131.5</td>
<td>453.0</td>
<td>129.8</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>290.0</td>
<td>335.6</td>
<td>461.0</td>
<td>334.4</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>337.0</td>
<td>960.7</td>
<td>461.0</td>
<td>919.4</td>
<td>/</td>
</tr>
</tbody>
</table>
Table 4 - 7: Tabulated excitation/emission data for 6 μM solutions of $^\bullet$dG mTBA.

<table>
<thead>
<tr>
<th>Position</th>
<th>Strand</th>
<th>$\lambda_{ex}$ (nm)</th>
<th>Intensity</th>
<th>$\lambda_{em}$ (nm)</th>
<th>Intensity</th>
<th>$I_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Na$^+$ Quad</td>
<td>260.0</td>
<td>18.8</td>
<td>380.0</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>471.9</td>
<td>18.9</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>325.1</td>
<td>7.4</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>473.0</td>
<td>6.5</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quad</td>
<td>255.0</td>
<td>5.6</td>
<td>378.0</td>
<td>5.1</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>460.0</td>
<td>2.7</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>315.1</td>
<td>2.8</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>460.0</td>
<td>3.3</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>260.0</td>
<td>5.2</td>
<td>371.1</td>
<td>1.5</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>483.9</td>
<td>5.3</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>325.1</td>
<td>5.1</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>467.0</td>
<td>4.6</td>
<td>/</td>
</tr>
<tr>
<td>8</td>
<td>K$^+$ Quad</td>
<td>262.0</td>
<td>21.0</td>
<td>494.0</td>
<td>23.5</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>268.0</td>
<td>38.7</td>
<td>480.0</td>
<td>38.5</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>327.1</td>
<td>39.1</td>
<td>/</td>
</tr>
</tbody>
</table>

Table 4 - 8: Tabulated excitation/emission data for 6 μM solutions of $^\bullet$y dG mTBA.

<table>
<thead>
<tr>
<th>Position</th>
<th>Strand</th>
<th>$\lambda_{ex}$ (nm)</th>
<th>Intensity</th>
<th>$\lambda_{em}$ (nm)</th>
<th>Intensity</th>
<th>$I_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>K$^+$ Quad</td>
<td>240.0</td>
<td>870.2</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>264.0</td>
<td>797.0</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>357.1</td>
<td>598.6</td>
<td>463.9</td>
<td>607.4</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>277.0</td>
<td>460.8</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>377.0</td>
<td>702.1</td>
<td>458.0</td>
<td>674.0</td>
<td>/</td>
</tr>
<tr>
<td>8</td>
<td>K$^+$ Quad</td>
<td>243.0</td>
<td>925.7</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>278.0</td>
<td>753.6</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>361.1</td>
<td>561.9</td>
<td>457.0</td>
<td>569.5</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>282.0</td>
<td>494.3</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>368.0</td>
<td>739.2</td>
<td>458.0</td>
<td>732.4</td>
<td>/</td>
</tr>
</tbody>
</table>
Some profound spectral changes were observed depending on whether TBA was folded into its quadruplex or corresponding duplex. All 8-aryl-dG adducts within quadruplex structures display a pronounced peak in the excitation spectrum at approximately 290 nm (except O\textsubscript{d}G and Py\textsubscript{d}G, which show a shoulder). Typically, the excitation spectra of 8-substituted guanosines exhibit two maxima corresponding to direct excitation (300-400 nm) and indirect excitation via unmodified nucleobases (260 nm).\textsuperscript{63} Clearly, the appearance of a third excitation maximum centred at 290 nm indicates some indirect excitation via the quadruplex, which, to our knowledge has not been documented previously. Interestingly, this excitation maximum was not present in 3+1 parallel-anti-parallel hybrid quadruplexes that were modified with C8-pyridyl-dG, C8-phenylethenyl-dG or C8-pyridyl-ethenyl-dG residues.\textsuperscript{63,64} Therefore, it seems as though the presence of this excitation peak at 290 nm is adduct- and situation-dependent, and may only arise when the proper modification is present in a heteropolar stacking arrangement found within an anti-parallel quadruplex.

Apart from the physical origin of this peak at 290 nm, its presence is clearly diagnostic for distinguishing quadruplex from duplex structures, as none of the studied duplexes have displayed this signal. This prominent peak affords much larger I\textsubscript{rel} values from excitation at 290 nm compared to exciting the adducts at their corresponding \(\lambda\text{max}\) values since the duplexes intensities at 290 nm are always very low. With all adducts, the fluorescence signal resulting from excitation at 290 nm increases at least two-fold when the quadruplex is formed (Table 4 - 4 - Table 4 - 8).

As can be seen in Figure 4 - 10 and Table 4 - 4, Fur\textsubscript{d}G was much more emissive (I\textsubscript{rel} ≥ 6) within a quadruplex than within a duplex. It also exhibited positional dependence, as I\textsubscript{rel} increased in the order of positions 5 < 8 < 6. It is important to note that Fur\textsubscript{d}G has previously been deemed an insensitive probe for use in duplex environments; we now find that it functions very well within a quadruplex environment. The Fur\textsubscript{d}G modification was the only adduct that was tolerated in a Na\textsuperscript{+}-stabilized quadruplex at position 5; the probes response was very similar to the
K+ stabilized quadruplex.

Although the BthdG generated greater emission intensities within the quadruplex environment (intensities were generally larger than the other 8 aryl dG probes studied), the probe was also very emissive in a duplex environment. This limits the sensitivity for distinguishing the quadruplex vs. duplex environment; this can be seen in Table 4 - 5, as BthdG exhibits much smaller I_rel values (I_rel ≤ 3). The BthdG emission intensity was somewhat sensitive to the position of modification within the quadruplex. BthdG@8 was quenched compared to positions 5 and 6. The BthdG at position 8 is stacking with the neighbouring tetrad, which has resulted in quenching of its emission.

CNPhdG shows very interesting changes in relative emission intensities. When position 5 of TBA is considered, the emission intensity of CNPhdG modified oligonucleotides is very similar for duplex and quadruplex structures (I_rel = 1.1). Position 6 is markedly better (I_rel = 1.8). However, CNPhdG shows significant positional dependence as modification at position 8 results in nearly a 10 fold difference in emission intensity when comparing quadruplex vs. duplex (I_rel = 0.1). This is due to the modified duplex being very emissive. In line with these data, the QdG also displays similar trends in I_rel, although not as pronounced (I_rel = 0.3). The emission of the D-A derivatives is clearly sensitive to the neighbouring bases, as modification of position 8 which is flanked by two Ts is the most emissive in all duplexes studied (including mNarI oligonucleotides). This parallels studies that report significant quenching for 2-AP and other fluorescent guanine analogues that were in close proximity and participated in base-stacking interactions with purines, especially guanosines.64,65 Mechanistically, such quenching has been attributed to photoinduced electron transfer (PET). Considering the fact that dG has the lowest oxidation potential of the nucleobases, it seems reasonable to suggest that the adducts’ emission is more effectively quenched due to more efficient electron transfer from guanosine in positions 5 and 6 than position 8. Assembly of the quadruplex brings one face of the adduct into direct contact with the G-tetrad which results in quenched emission intensities.
Interestingly, the PET quenching is much more pronounced for the D-A “push-pull” 8-aryl-dG nucleosides and is responsible for quenching their CT emissions. The direction of electron transfer in the excited state is determined by the oxidation and reduction potentials of the ground and excited states. All 8-aryl-dG modified nucleosides have lower oxidation potentials than the unmodified dG nucleoside; therefore, based on the oxidation potentials, one would assume the electron transfer would proceed from the 8-aryl-dG adduct to the dG nucleoside. However, although this may happen to a small extent, it clearly is not the dominant mechanism for PET quenching as all 8-aryl-dG analogues would experience similar levels of quenching, and one would then expect larger amounts of quenching in the presence of the less electron-dense dA, dC or dT nucleosides (more able to accept electrons). An alternative explanation may be that, upon excitation the D-A nucleosides generate charge-separated biradicaloid species produced by electron transfer (ET) from the donor to acceptor when the two aryl rings are highly twisted from planarity. Thus, in the present case, the CNPh\textsubscript{dG} emission at 462 nm is ascribed to a CT-state resulting from ET to afford a G*+ CNPh*− species. When this CT state is generated CNPh\textsubscript{dG} is more prone to quenching via electron transfer from a neighbouring G (the most electron dense nucleobase) into the G*+ component of CNPh\textsubscript{dG} as it is electron deficient.

Although Q\textsubscript{dG} was the least emissive of the adducts, modification of TBA at position 5 resulted in dual fluorescence at 390 and 480 nm respectively. In the K\textsuperscript{+}-stabilized quadruplex the Q\textsubscript{dG} displayed emission intensity at 390 nm with a small shoulder at 480 nm while in the Na\textsuperscript{+}-stabilized quadruplex Q\textsubscript{dG} was more emissive at 480 nm. It is expected that in either the K\textsuperscript{+} or Na\textsuperscript{+} quadruplex the Q\textsubscript{dG} is projected into the aqueous environment, which typically quenches the 480 nm CT-emission. The specific molecular interactions that are occurring to result in the changes of emission for Q\textsubscript{dG} are unknown.

Py\textsubscript{dG} responded similar to Bth\textsubscript{dG}, the adduct had large emission intensities in the quadruplex but was equally emissive in the duplex environment (I\textsubscript{rel} ≈ 1); therefore, Py\textsubscript{dG} was not able to distinguish between the two.
Modifying two positions within TBA was aimed at designing more effective oligonucleotide systems for distinguishing duplex vs. quadruplex structures. Clearly, the $T_m$ values of the doubly modified oligonucleotides significantly destabilize the duplex while the modifications seem to be well tolerated by the quadruplexes (Table 4-3). This allows for quadruplex formation to effectively compete with the duplex; the fluorescent responses of these modifications are shown below in Figure 4-11.

**Figure 4-11**: Excitation and emission spectra of (a) $\text{Fur} \cdot \text{dG}_{5/8}$, (b) $\text{Fur} \cdot \text{dG}_{10/8}$, (c) $\text{Fur} \cdot \text{dG}_{1/8}$, and (d) $\text{Fur} \cdot \text{dG}_{5/8}$ mTBA oligonucleotides. Dotted line represents duplex. Dashed line represents Na$^+$ quadruplex. Solid Line represents K$^+$ quadruplex.
Table 4 - 9: Tabulated excitation and emission spectra of the dual modified mTBA oligonucleotides.

<table>
<thead>
<tr>
<th>mTBA</th>
<th>Structure</th>
<th>$\lambda_{ex}$ (nm)</th>
<th>Intensity</th>
<th>$\lambda_{em}$ (nm)</th>
<th>Intensity</th>
<th>$I_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Fur dG}_{5}, \text{CNPh dG@8}$</td>
<td>Na’ Quad</td>
<td>325.1</td>
<td>229.6</td>
<td>460.0</td>
<td>224.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>K’ Quad</td>
<td>325.1</td>
<td>116.2</td>
<td>456.1</td>
<td>108.5</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>340.0</td>
<td>753.5</td>
<td>460.0</td>
<td>689.9</td>
<td>/</td>
</tr>
<tr>
<td>$\text{Fur dG@10}, \text{CNPh dG@5}$</td>
<td>Na’ Quad</td>
<td>325.1</td>
<td>213.5</td>
<td>458.0</td>
<td>209.9</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>K’ Quad</td>
<td>232.1</td>
<td>151.3</td>
<td>458.9</td>
<td>163.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>337.0</td>
<td>499.5</td>
<td>460</td>
<td>475.5</td>
<td>/</td>
</tr>
<tr>
<td>$\text{Fur dG@1}, \text{Py dG@8}$</td>
<td>K’ Quad</td>
<td>356.0</td>
<td>61.1</td>
<td>461.9</td>
<td>62.8</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>372.0</td>
<td>739.2</td>
<td>458.9</td>
<td>737.1</td>
<td>/</td>
</tr>
<tr>
<td>$\text{Fur dG@5}, \text{Py dG@8}$</td>
<td>K’ Quad</td>
<td>360.0</td>
<td>85.7</td>
<td>461.0</td>
<td>89.3</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>370.0</td>
<td>702.3</td>
<td>458.0</td>
<td>713.2</td>
<td>/</td>
</tr>
</tbody>
</table>

In the $\text{Fur dG/\text{CNPh dG}}$ modified oligonucleotide, the goal of the design was to exploit the large differences in emission intensity for $\text{CNPh dG@8}$, while allowing the $\text{Fur dG}$ modification to stabilize the overall quadruplex structure. As can be seen in Figure 4 - 11, the oligonucleotide functioned as desired, as it was clearly able to distinguish between quadruplex and duplex structures. The $\text{Fur dG/Py dG}$ also enjoyed success at distinguishing between duplex and quadruplex, showing significantly quenched emission spectra upon formation of the quadruplex.

4.4.3. Titrations

4.4.3.1. K$^+$ Titrations

To demonstrate the potential utility of the modified TBA samples for DNA-based diagnostics, duplex-quadruplex exchange studies were performed using TBA with $\text{Fur dG@5}$ and the doubly-incorporated sample with $\text{Fur dG@5, CNPh dG@8}$. These modified TBA samples have $T_m$ values ≈ 60°C for the G-quadruplex in K$^+$, that are approximately 4.5-11°C higher than $T_m$ values for the fully-complementary duplex in Na’. In order for K$^+$ to induce duplex unwinding and generate the G-quadruplex structures, the mTBA samples were initially annealed to their
complementary sequence (either a 10-mer sequence 5′-CCACACCAAC or full complement). Truncated sequences were utilized to lower duplex stability in 100 mM Na+-phosphate buffer pH 7.0 with 0.1 M NaCl. With this 10-mer sequence the Fur@5 mTBA oligonucleotide had a duplex $T_m = 41^\circ$C, which is 21.4$^\circ$C below the $T_m$ value for the G-quadruplex Fur@5 mTBA in K+. For the doubly-incorporated TBA, the full-length 15-mer duplex was employed, which provided a duplex with $T_m = 49.4$ $^\circ$C, that is 10.6 $^\circ$C below the $T_m$ of the G-quadruplex. Titrations of the duplexes with KCl are depicted in Figure 4 - 12. For Fur@5 mTBA the duplex with the 10-mer complementary strand exhibited quenched fluorescence as expected (Figure 4 - 12) and provided a CD spectrum consistent with B-form DNA (Figure 4 - 12, insert, solid red trace). Additions of KCl (5 mM aliquots) caused the fluorescence intensity to increase with growth in the diagnostic 290 nm energy-transfer band in the excitation spectrum for G-quadruplex formation (dotted red traces, Figure 4 - 12). Under conditions of the KCl titration (25$^\circ$C), the conversion of the duplex into the G-quadruplex structure was relatively slow, and the equilibrium was monitored over a 2 day period (6h between each addition of KCl). After the final addition of KCl, a CD spectrum of the mixture confirmed formation of the anti-parallel G-quadruplex (Figure 4 - 12 insert, dotted red trace).
Figure 4 - 12: Probe fluorescence response in duplex–quadruplex exchange mediated by K⁺. Duplex samples represented by solid trace. KCl was then added in 5 mM aliquots (dotted traces). Inserts show CD overlays of initial duplex (solid trace) and G-quadruplex after the final KCl addition (dotted trace). (a) Titration of Fur dG@5 mTBA:10-mer with KCl at 25 °C, (b) titration of Fur dG@5/CNPh dG@8 mTBA:15-mer with KCl at 30°C.

The corresponding KCl titration of the duplex containing Fur dG@5/CNPh dG@8 mTBA is shown in Figure 4 - 12 (b). In this case, the duplex is strongly emissive at 450 nm due to the CNPh dG probe at G₈. Addition of KCl to the duplex at 15°C caused quenching of the CNPh dG emission at 450 nm (dotted blue traces). This observation was consistent with G-quadruplex formation, as folding into the G-quadruplex quenches the emission of CNPh dG due to stacking interactions with the G-tetrad. The emission response of the doubly-incorporated probe was dominated by the emission of CNPh dG and a separate emission peak at ~ 380 nm for Fur dG at G₅ was not detected. Thus, the presence of the D-A CNPh dG probe at G₈ inhibited energy-transfer from the G-tetrad into Fur dG, and the Fur dG probe at G₅ served to increase G-quadruplex stability and decrease duplex stability, while remaining optically silent during the duplex-quadruplex exchange. It is also important to note that compared with the titration of Fur dG@5 mTBA:10-mer with KCl, the conversion of the full-length duplex into the G-quadruplex structure for Fur dG@5/CNPh dG@8 mTBA required larger amounts of KCl. Once the titration was deemed complete the CD spectrum confirmed G-quadruplex formation (insert Figure 4 - 12, dotted blue...
trace). The difference in relative amounts of added KCl was ascribed to the relative duplex stabilities, as revealed by the DNA melting experiments ($T_m$ 41 vs. 49.4 °C).

4.4.3.2. Thrombin Binding Titrations

Aside from determining the stability of the modified TBA sequences and the ability of these sequences to monitor duplex to quadruplex exchange reactions, it was important to determine how the modified TBA sequences impact the aptamer’s ability to recognize its molecular target (thrombin).

Thrombin binding was probed in the following way: 1 mL mTBA solutions were prepared (6 μM) in 100 mM sodium phosphate buffer pH 7.0 with 0.1 M NaCl. This solution was titrated with aliquots (0.4 μM) of thrombin at 20 min intervals and monitored by fluorescence. Titrations were carried out until no further changes in fluorescence were detected upon additions of thrombin. Alternatively, some samples were designed to undergo duplex-quadruplex exchange, driven by thrombin. For these titrations, duplexes (6 μM) were prepared by mixing mTBA with its complementary sequence (5'-CCAACCACACCAACC) in the aforementioned buffer prior to titrations. Representative titrations are shown from Figure 4 - 13 to Figure 4 - 25 while the association constants ($K_a$) for thrombin binding by the mTBA oligonucleotides are presented in Table 4 - 10. The stoichiometry of the binding reaction indicated a 1:1 aptamer/thrombin interaction. In general, $K_a$ values are approximately 1 order of magnitude below previously reported values for thrombin binding by unmodified TBA in K$^+$ buffer. This correlates with the lower stability of the thrombin-TBA-Na$^+$ compared to that of thrombin-TBA-K$^+$. The probes are capable of signalling thrombin binding. Specificity was confirmed by control titrations with bovine serum albumin (BSA) (Figure 4 - 26 and Figure 4 - 27), which resulted in less than 1% change in emission intensity.
Figure 4 - 13: ssDNA to quadruplex fluorescence titration of TBA F^4dG@5 (λ_{ex} = 320 nm, λ_{em} 380 nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

\[ K_d = (3.33 \pm 0.41) \times 10^{-6} \]
\[ R^2 = 0.960 \]

Figure 4 - 14: ssDNA to quadruplex fluorescence titration of TBA F^4dG@6 with thrombin (λ_{ex} = 320 nm, λ_{em} 380 nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

\[ K_d = (5.23 \pm 0.88) \times 10^{-6} \]
\[ R^2 = 0.946 \]
Figure 4 - 15: ssDNA to quadruplex fluorescence titration of TBA $^{Fur}$dG@8 with thrombin ($\lambda_{ex} = 320$ nm, $\lambda_{em} 380$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

Figure 4 - 16: ssDNA to quadruplex fluorescence titration of TBA $^{Bth}$dG@5 with thrombin ($\lambda_{ex} = 330$ nm, $\lambda_{em} 420$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.
Figure 4 - 17: ssDNA to quadruplex fluorescence titration of TBA $^{2}$dG@8 with thrombin ($\lambda_{ex} = 330$ nm, $\lambda_{em} 420$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

$$K_d = (8.90 \pm 0.97) \times 10^{-6}$$

$$R^2 = 0.965$$

Figure 4 - 18: ssDNA to quadruplex fluorescence titration of TBA $^{CNP}$dG@5 with thrombin ($\lambda_{ex} = 330$ nm, $\lambda_{em} 455$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

$$K_d = (6.43 \pm 0.98) \times 10^{-6}$$

$$R^2 = 0.961$$
Figure 4 - 19: dsDNA to quadruplex fluorescence titration of TBA\textsuperscript{CNPh}dG@5 with thrombin ($\lambda_{\text{ex}} = 330$ nm, $\lambda_{\text{em}} 455$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

![Graph of dsDNA to quadruplex fluorescence titration of TBA\textsuperscript{CNPh}dG@5 with thrombin (λ\text{ex} = 330 nm, λ\text{em} 455 nm).](image1)

$K_d = (9.42\pm1.23) \times 10^{-6}$

$R^2 = 0.953$

Figure 4 - 20: ssDNA to quadruplex fluorescence titration of TBA\textsuperscript{Fur}dG@5/CNPh\textsuperscript{CNPh}dG@8 with thrombin ($\lambda_{\text{ex}} = 320$ nm, $\lambda_{\text{em}} 455$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

![Graph of ssDNA to quadruplex fluorescence titration of TBA\textsuperscript{Fur}dG@5/CNPh\textsuperscript{CNPh}dG@8 with thrombin (λ\text{ex} = 320 nm, λ\text{em} 455 nm).](image2)

$K_d = (3.82\pm0.43) \times 10^{-6}$

$R^2 = 0.967$
Figure 4 - 21: ssDNA to quadruplex fluorescence titration of TBA $^{\text{Fur}}$dG@1/$^{\text{fly}}$dG@8 with thrombin ($\lambda_{\text{ex}} = 350$ nm, $\lambda_{\text{em}} 455$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

$$K_d = \left(5.54 \pm 0.96\right) \times 10^{-6}$$
$$R^2 = 0.940$$

Figure 4 - 22: dsDNA to quadruplex fluorescence titration of TBA $^{\text{Fur}}$dG@5 with thrombin ($\lambda_{\text{ex}} = 320$ nm, $\lambda_{\text{em}} 380$ nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.
Figure 4 - 23: dsDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@5/CNPdG@8 with thrombin ($\lambda_{\text{ex}} = 330$ nm, $\lambda_{\text{em}}$ 455 nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

\begin{align*}
\text{Fraction Bound} & \quad 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \quad 1.2 \\
\text{[Thrombin] (M)} & \quad 0.0 \quad 5.0\times10^{-6} \quad 1.0\times10^{-5} \quad 1.5\times10^{-5} \quad 2.0\times10^{-5} \quad 2.5\times10^{-5} \\
K_d & = (4.41\pm0.79) \times 10^{-6} \\
R^2 & = 0.933
\end{align*}

Figure 4 - 24: dsDNA to quadruplex fluorescence titration of TBA \textsuperscript{Fur}dG@1/PydG@8 with thrombin. Here excitation and emission wavelengths were set to monitor the \textsuperscript{Py}dG moiety ($\lambda_{\text{ex}}$ 360 nm, $\lambda_{\text{em}}$ 460 nm). Insert: Binding isotherm plotting Fraction of oligonucleotide bound vs. [thrombin] to generate dissociation constants and associated coefficient of determination.

\begin{align*}
\text{Fraction Bound} & \quad 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \quad 1.2 \\
\text{[Thrombin] (M)} & \quad 0.0 \quad 5.0\times10^{-6} \quad 1.0\times10^{-5} \quad 1.5\times10^{-5} \quad 2.0\times10^{-5} \quad 2.5\times10^{-5} \\
K_d & = (1.06\pm0.21) \times 10^{-6} \\
R^2 & = 0.920
\end{align*}
Figure 4 - 25: dsDNA to quadruplex fluorescence titration of TBA $^{\text{Fur}}$dG@1/$^{\text{Py}}$dG@8. Here excitation and emission wavelengths were set to monitor the $^{\text{Fur}}$dG probe ($\lambda_{\text{ex}}$ 315 nm, $\lambda_{\text{em}}$ 380 nm).

Table 4 - 10: Tabulated association constants and corresponding coefficient of determination values for thrombin binding by 8-aryl-dG mTBA oligonucleotides

<table>
<thead>
<tr>
<th>Starting Structure$^a$</th>
<th>Adduct</th>
<th>$K_d$ (x10$^{-6}$)$^b$</th>
<th>$K_a$ (x10$^5$)</th>
<th>$R^2$</th>
<th>$I_{\text{rel}}$$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dsDNA</td>
<td>$^{\text{Fur}}$dG@1; $^{\text{Py}}$dG@8</td>
<td>1.06</td>
<td>9.43</td>
<td>0.920</td>
<td>0.82</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Fur}}$dG@1; $^{\text{Py}}$dG@8</td>
<td>2.32</td>
<td>4.31</td>
<td>0.901</td>
<td>0.70</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Fur}}$dG@5</td>
<td>3.33</td>
<td>3.00</td>
<td>0.960</td>
<td>2.07</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Bth}}$dG@5</td>
<td>3.80</td>
<td>2.63</td>
<td>0.968</td>
<td>1.31</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Fur}}$dG@5; $^{\text{CNPh}}$dG@8</td>
<td>3.82</td>
<td>2.62</td>
<td>0.967</td>
<td>0.50</td>
</tr>
<tr>
<td>dsDNA</td>
<td>$^{\text{Fur}}$dG@5; $^{\text{CNPh}}$dG@8</td>
<td>4.41</td>
<td>2.27</td>
<td>0.933</td>
<td>0.28</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Fur}}$dG@8</td>
<td>4.50</td>
<td>2.22</td>
<td>0.979</td>
<td>1.41</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Fur}}$dG@6</td>
<td>5.23</td>
<td>1.91</td>
<td>0.946</td>
<td>1.26</td>
</tr>
<tr>
<td>dsDNA</td>
<td>$^{\text{Fur}}$dG@5</td>
<td>5.54</td>
<td>1.81</td>
<td>0.940</td>
<td>7.43</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{CNPh}}$dG@5</td>
<td>6.43</td>
<td>1.56</td>
<td>0.961</td>
<td>0.25</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{Bth}}$dG@8</td>
<td>8.90</td>
<td>1.12</td>
<td>0.965</td>
<td>1.19</td>
</tr>
<tr>
<td>ssDNA</td>
<td>$^{\text{CNPh}}$dG@8</td>
<td>9.42</td>
<td>1.06</td>
<td>0.953</td>
<td>0.40</td>
</tr>
</tbody>
</table>

$^a$ Starting structure refers to whether the titration was performed on the single stranded (ssDNA) or preformed double stranded (dsDNA) oligonucleotides.  
$^b$ All determined dissociation constants were obtained through SigmaPlot simple ligand binding macro.  
$^c$ $I_{\text{rel}}$ calculated as fluorescence intensity of the starting oligonucleotide divided by fluorescence intensity of the fully bound TBA.
Clearly, the probes are capable of depicting thrombin binding. Unfortunately, $^{B,H}_{d}G@6$, $^{CNPh}_{d}G@6$, and all $^{Q}_{d}G$ modifications were not able to provide reliable diagnostic signals in the thrombin binding titrations. The data in Table 4 - 10 are arranged in descending order with regard to association constant. Association constants were determined using the “simple ligand
binding” macro within SigmaPlot version 11.0, and are in relatively good agreement with a 1:1 binding stoichiometry, as \( R^2 \) values range from 0.901 to 0.979.

Interestingly, the oligonucleotide that generated the largest two association constants for thrombin binding were the doubly modified \( \text{Fur}dG@1/\text{Py}dG@8 \) mTBA oligonucleotides; interestingly, the dsDNA to quadruplex transition providing a larger association constant than the ssDNA to quadruplex transition. This was unexpected, the doubly modified oligonucleotides displayed significantly less stable quadruplex structures (\( \Delta T_m = -7.7^\circ C \)) compared to the other single modifications at position 5 (1.9\(^\circ C \leq \Delta T_m \leq 9.1\(^\circ C \)); lower melting temperatures were anticipated to result in smaller association constants. These results may be accounted for by previous observations by Krauss et al. who report that interactions between TBA and \( \alpha \)-thrombin are both hydrophobic and hydrophilic, and involve residues of the TT loops with a further contribution of G5.\(^{44} \) Modification of G5 with C8-aryl-dG residues stabilize the overall quadruplex structure; however, they seem to interfere with the aptamer’s binding to thrombin, possibly due to the steric bulk added by the modification. In these systems, fluorescence emission is quenched as thrombin is bound. The \( \text{Fur}dG@5/\text{Py}dG@8 \) dsDNA to quadruplex transition provided smaller fluorescent responses (\( I_{rel} = 0.82 \)) than the ssDNA to quadruplex scenario (\( I_{rel} = 0.70 \)).

Probe performance based on location within the aptamer is well depicted within the \( \text{Fur}dG \) series of modifications. Overall the association constant correlates with the stability of the quadruplex structure and increases from \( \text{Fur}dG@6 (K_a = 1.91 \times 10^5; T_m = 42.3^\circ C) < \text{Fur}dG@8 (K_a = 2.22 \times 10^5; T_m = 43.8^\circ C) < \text{Fur}dG@5 (K_a = 3.00 \times 10^5; T_m = 62.4^\circ C) \). All \( \text{Fur}dG \) probes display enhanced fluorescence emission upon transitioning from either ssDNA or dsDNA to quadruplex. \( \text{Fur}dG@5 \) provided the largest \( I_{rel} \) for the series of modifications (\( I_{rel} = 2.07 \)) followed by \( \text{Fur}dG@8 \) (\( I_{rel} = 1.41 \)) and \( \text{Fur}dG@6 \) (\( I_{rel} = 1.26 \)). Fluorescence response was further improved for \( \text{Fur}dG@5 \) by performing a duplex to quadruplex transition titration, in which \( \text{Fur}dG@5 \) was hybridized to its
complementary sequence to form the corresponding duplex prior to introduction of thrombin. This resulted in larger relative intensities ($I_{rel} = 7.43$), but also lowered the overall association constant ($K_a = 3.00 \times 10^5$ for the ssDNA to quadruplex titration as opposed to $1.81 \times 10^5$ for the dsDNA to quadruplex titration). This wasn’t the case for the doubly modified $\text{Fur}dG@1/Py\text{G}@8$ duplex; it is likely that the large destabilization ($\Delta T_m = -22^\circ C$) in duplex structure imposed by the dual $Py\text{dG}/\text{Fur}\text{dG}$ modifications resulted in little resistance for duplex unwinding and quadruplex formation. The $\text{Fur}dG@5$ duplex was clearly not as destabilized ($\Delta T_m = -6.6$), and the duplex structure was more able to compete with quadruplex formation throughout the titration.

$\text{Bth}dG$ probes were tested for only the ssDNA to quadruplex transitions. This is due to the $\text{Bth}dG$ modifications having similar fluorescent intensities in duplex vs. quadruplex environments, which limits their use for detecting DNA transitions in situations where excitation at 290 nm is not possible (thrombin $\lambda_{max} = 280$ nm). $\text{Bth}dG$ probes have slightly larger emission intensities in the quadruplex than in ssDNA ($I_{rel} < 1.3$). The thrombin-induced ssDNA to quadruplex titrations generated association constants that again were clearly correlated to overall quadruplex stability, as $\text{Bth}dG@5$ ($K_a = 2.63 \times 10^5$; $T_m = 55.2^\circ C$) was greater than $\text{Bth}dG@8$ ($K_a = 1.12 \times 10^5$; $T_m = 45.2^\circ C$). Overall, the $\text{Bth}dG$ modification was not as effective at probing thrombin binding as $\text{Fur}dG$; its association constants were lower, as was its fluorescence response to thrombin binding and quadruplex formation.

Unlike $\text{Fur}dG$ and $\text{Bth}dG$ probes, the $\text{CNPh}dG$ probes’ fluorescence emission was quenched upon thrombin binding, whether the transition was from ssDNA or dsDNA to quadruplex. The response of the $\text{CNPh}dG$ probes were quite good, as $\text{CNPh}dG@5$ quenched 4 fold upon thrombin binding ($I_{rel} = 0.25$), while $\text{CNPh}dG@8$ was quenched approximately 2.5 fold ($I_{rel} = 0.4$). Among the modifications studied, the association constant was again related to overall quadruplex stability as $\text{CNPh}dG@5$ ($K_a = 1.56 \times 10^5$; $T_m = 60.4^\circ C$) was larger than $\text{CNPh}dG@8$ ($K_a = 1.06 \times 10^5$; $T_m = 47^\circ C$). The relatively low association constant for $\text{CNPh}dG@5$ was not expected, as the
quadruplex was more stable than the $^{\text{Bth}}$dG@5. Sterics alone cannot explain the differences, as the $^{\text{Bth}}$dG would be expected to have a more detrimental effect in that regard. The major difference between $^{\text{CNPh}}$dG and $^{\text{Fur}}$dG or $^{\text{Bth}}$dG is that the $^{\text{CNPh}}$dG is a donor-acceptor probe. $^{\text{CNPh}}$dG exhibits the largest twist angles in the ground state ($\theta = 37.9^\circ$) as opposed to $^{\text{Fur}}$dG ($\theta = 16.2^\circ$) or $^{\text{Bth}}$dG ($\theta = 31.1^\circ$). This could result in unfavourable interactions between $^{\text{CNPh}}$dG@5 with thrombin which would result in lower association constants.

In an attempt to improve the association constants with the $^{\text{CNPh}}$dG modification, which was the only probe to provide quenched fluorescence upon thrombin binding, the doubly modified $^{\text{Fur}}$dG@5/$^{\text{CNPh}}$dG@8 oligonucleotide was subjected to thrombin titrations. Adding the $^{\text{Fur}}$dG modification at position 5 improved the interaction of the mTBA with thrombin as the ssDNA to quadruplex titration ($K_a = 2.62 \times 10^5$) and the dsDNA to quadruplex titration ($K_a = 2.27 \times 10^5$) for doubly modified TBA were larger than that of the singly incorporated $^{\text{CNPh}}$dG@8 ($K_a = 1.06 \times 10^5$). The dsDNA to quadruplex transition also exhibited larger changes in fluorescence quenching ($I_{\text{rel}} = 0.28$) than the ssDNA to quadruplex transition ($I_{\text{rel}} = 0.5$).

4.5. Conclusions

The potential utility of internal fluorescent 8-aryl-dG probes for monitoring duplex-quadruplex exchange, a common strategy in DNA-based diagnostics, was demonstrated. The probes adopt a syn-conformation and are accommodated within the loop-positions of the anti-parallel G-quadruplex formed by the 15-mer thrombin-binding aptamer (TBA). Within the G-tetrad, the probes stabilize an anti-parallel G-quadruplex when placed in a syn-G position, but are strongly destabilizing in an anti-G position. Our studies show that aryl selection can be used to manipulate the emissive response of the 8-aryl-dG probe to duplex-quadruplex exchange. The $^{\text{Fur}}$dG probe exhibits quenched emission at 380 nm ($\lambda_{\text{ex}} = 310$ nm) in duplex DNA that turns-on in the G-quadruplex due to energy-transfer from unmodified dG within the tetrad. In fact, the $^{\text{Fur}}$dG probe provides a diagnostic excitation spectrum at 290 nm for G-quadruplex folding. In contrast,
the push-pull donor-acceptor $^{\text{CNPh}}_{\text{dG}}$ probe is strongly emissive at 460 nm ($\lambda_{\text{ex}} = 330$ nm) in the duplex that turns-off in the G-quadruplex. In this case, the emission of the D-A probe at 460 nm is ascribed to a charge-transfer state possessing $\text{G}^{\text{+}}\text{CNPh}^{-}$ character due to electron-transfer (ET) from the donor (dG) into the acceptor (CNPh). The quenching of $^{\text{CNPh}}_{\text{dG}}$ emission upon G-quadruplex folding was ascribed to stacking interactions with G residues in the tetrad leading to ET from the electron-rich G into $\text{G}^{\text{+}}$ of the charge-separated species.

Utilizing 8-aryl-dG modifications that can stabilize G-quadruplexes and provide diagnostic signals upon quadruplex folding provide numerous applications for aptamer development. G-Tetrad positions are much less variable than terminus, or loop regions of G-quadruplex folding aptamers. Incorporation of internal fluorescent G-mimics within G-tetrads can be advantageous for investigations of quadruplex folding where the level of detail needs to be high since the probe can be designed to be very close to the site of examination. The most important implication these studies have presented involves the utilization of 8-aryl-dG nucleobase analogues within G-tetrads provides a probing scenario that should be universally applicable to G-quadruplex aptamers folding into anti-parallel topologies.
4.6. References


(49) Smirnov, I.; Shafer, R. H. Biochemistry 2000, 39, 1462.


5.1. BIG PICTURE

Over 50 years ago Leonard Hayflick and his colleague Paul Moorhead discovered that cultured normal human cells have a limited number of generations to divide, after which they stop growing, become enlarged, and engage in a new pathway termed replicative senescence. At the time this was totally unexpected as the research community was under the impression that cells explanted into cell culture were immortal. Hayflick and Moorhead also demonstrated that cells had a molecular counting mechanism regardless of their surrounding cells. They demonstrated that this counting mechanism was somehow programmed into each cell, and once this biological clock had expired, the cell would stop dividing. Further studies established that the transition from proliferation to senescence was controlled by telomere length. Studies on replicative senescence have begun to elucidate aspects of tissue and organismal aging, and have created new opportunities in the areas of regenerative medicine for aging tissues and telomeropathies, that is, genetic diseases due to premature telomere shortening. Cancer cells have the ability to overcome senescence, by utilizing mechanisms that enable the maintenance of telomere length; this enables cancer cells to divide indefinitely, and has become a biomarker of almost all advanced human cancers.

In this chapter the previously introduced FurG probe will be utilized to analyze various folding topologies within the human telomeric sequence. The sequence is highly polymorphic and has been shown to fold into various G-quadruplexes. Interestingly, there seems to be major biological implications dependent on which type of G-quadruplex forms. The FurG probes have previously been shown to distinguish major differences in DNA tertiary structure (i.e. single strand vs. duplex vs. quadruplex). This chapter will focus on the ability for FurG to distinguish between the different types of G-quadruplexes, which would diversify the probe’s applicability. Ideally, this would allow it to gain use for in vitro tests to study telomeres, and determine this sequences interactions with small molecules or proteins which induce the formation of specific G-quadruplexes.
5.2. General Introduction

Telomeres are nucleoprotein complexes at the ends of linear chromosomes. In humans, they are composed of double stranded 5′-T\(_2\)AG\(_3\) repeats typically extending between 9-15 kb with a 3′ single-stranded overhang containing 50-300 nucleotides (termed a G-tail or G-overhang).\(^6\) In human somatic cells, the average bulk telomeric length is an inherited trait,\(^7\) but the replicative history of the cell as well as environmental factors including exposure to stress and oxidative damage are also key factors that affect telomere length.\(^8,9\)

Functionally, telomeres are implicated in the protection of chromosomes from degradation, end-to-end fusions, and from being recognized as double-strand break sites.\(^10\) Other essential functions of telomeres include (i) providing genetic stability with regard to the complete replication of the chromosome, (since many mammalian cells lose telomeric DNA at each division as opposed to coding genetic material); (ii) contributing to the spatial and functional organisation of chromosomes within the nucleus; and (iii) participating in regulation of transcription.\(^11\) Telomere length controls the replicative capacity of cells, implicating short telomere length in age-related diseases. Telomeres also appear to play additional, less defined roles in cells. For example, they have been implicated in meiosis\(^12,13\) and have been found to repress the expression of genes placed near them,\(^14\) although the natural role of telomere transcriptional silencing is unclear.

Telomeres require specialized replication machinery, as they are present at the end of chromosomes. The ends of linear DNA molecules cannot be replicated fully by the semi-conservative DNA replication machinery. This is termed the ‘end replication problem’ and is illustrated in Scheme 5 - 1. At the very end of the lagging strand, the RNA primer is not replaced with DNA, as there is no incoming 5′-3′ acting DNA polymerase δ to displace it. This RNA primer is then degraded, which regenerates the pre-existing 3′ G-overhang. The leading strand
remains blunt-ended and regeneration of the telomeric overhang by an exonuclease results in DNA loss.

Scheme 5 - 1: The end replication problem. DNA at the end of telomeres exists as a single-stranded 3' overhang. Parental duplex (black) is unwound at the replication fork and replication proceeds continuously on the leading strand and discontinuously on the lagging strand. RNA primers are shown in red while the generated DNA fragments are shown in blue.

Telomere maintenance is necessary for long term cell proliferation. It has been demonstrated that, in the absence of specific replication machinery at the telomere ends, gradual sequence loss due to incomplete replication of the lagging strand eventually leads to critically short telomeres and trimming of essential chromosomal sequences. In order to compensate for this loss, telomere maintenance is performed, mainly by a specific reverse transcriptase, telomerase. Human telomerase is a ribo-nucleoprotein composed of a catalytic subunit, hTERT and a 451 nt long RNA, which acts as a template for the addition of the short repetitive motif d(GGGTAA) on the 3' end of a primer. Telomerase is highly expressed in cells that need to divide regularly such as the immune system, germ line, as well as some stem cells. However, most somatic cells express it in very low levels, in a cell-cycle dependent manner. Interestingly,
telomerase is also up-regulated in most (over 85%) cancers, in which it is thought to promote the lifespan of malignant cells. Recent key experiments demonstrated that: (i) telomerase is sufficient for immortalisation of many cell types and (ii) inhibition of telomerase limits the growth of human cancer cells. All of these results point to a key role of telomerase in the tumorigenic process. Furthermore, unlimited proliferative potential, which depends on telomere maintenance, has been defined as one of the six hallmarks of cancer. Clearly, an understanding of telomere/telomerase regulation is expected to give major insights in the tumorigenesis process, and its manipulation is a challenge and goal for the design of future anti-cancer approaches.

5.2.1. Structure and Stability of Telomeres

As previously mentioned, a distinctive feature of telomeric DNA is that it terminates in a single stranded G-overhang. Typically, this single stranded DNA would be subject to nucleolytic attack and recombinational activities. The folding of this overhang into higher-order structures such as T-loops and G-quadruplexes provides the protection needed for genomic stability. In addition to such higher order structures, telomeric proteins coating the double strand-single strand boundaries provide protection. A series of six proteins have been identified in a single complex, which has become known as shelterin.

T-loops were first identified in the telomeric DNA of human and mouse cells by electron microscopy. The proposed function of T-loops is to protect telomeres and regulate telomerase access; however, it is unknown what proportion of human telomeres form T-loops or whether they are the predominant telomere capping structure. A schematic representation of a T-loop is displayed in Figure 5 - 1. T-loop formation seems to depend on the ability of certain components of the shelterin complex to remodel the linear telomere into the loop.
The alternative protective structures for the G-overhang of mammalian telomeres are G-quadruplexes. G-quadruplex DNA structures have attracted considerable attention, because they provide a simple mechanism for telomere-telomere interactions, and inhibit telomerase activity in vitro. Non-denaturing polyacrylamide gels run on oligonucleotides corresponding to the telomeric G-rich strands display unexpected banding patterns that are due to the formation of different G-quadruplex structures.

Aside from the potential capping and protective functions telomeric G-quadruplex structures provide, their role in other functions in human cells is relatively unknown. It is possible that different conformations are responsible for different functions. Intermolecular G-quadruplexes have been implicated in telomere-telomere associations, and in the alignment of sister chromatids during meiosis. Various hybrid intramolecular forms of the human telomeric sequence have the potential to stack end-to-end in long arrays, which may accomplish compaction of telomeric DNA. Many open questions remain.

While in vitro evidence of G-quadruplex DNA structures is ample, in vivo evidence is scarce. Through the use of structure-specific antibodies, direct evidence that an telomeric G-
quadruplex DNA structures are present \textit{in vivo} was reported in the macronucleus of stichotrichous ciliates\textsuperscript{30} and more recently in human cells.\textsuperscript{37} Intramolecular anti-parallel G-quadruplexes have been shown to be resistant to telomerase elongation \textit{in vivo},\textsuperscript{19,27,35} whereas parallel intermolecular structures are extended by telomerase.\textsuperscript{35} More \textit{in vivo} evidence stems from reports from Paeschke et al. which showed that telomere end-binding proteins (TEBP) have distinct, but essential roles in the formation and regulation of G-quadruplex DNA structures at telomeres.\textsuperscript{36} TEBPs seem to act as chaperones to promote G-quadruplex DNA formation. It remains to be seen how the G-quadruplex structures are unfolded \textit{in vivo} in a timely manner as they are very stable under physiological conditions.\textsuperscript{36}

\textbf{5.2.2. Structural Polymorphism of HTelo G-Quadruplexes}

Numerous studies have tried to determine which conformations are adopted by the G-quadruplex under physiological conditions.\textsuperscript{38-41} It appears evident that the quadruplexes of human telomeric DNA are quite heterogeneous, as the 5′-TTAGGG repetitive framework provides multiple opportunities for different quadruplexes to form under different experimental conditions.\textsuperscript{41-45} Multiple G-quadruplex conformations can be observed for a given sequence, which makes structural elucidation difficult.\textsuperscript{46} Researchers typically overcome such conformational heterogeneity by modification of the flanking nucleotides and/or base-analogue substitutions to favour one conformation over the other.\textsuperscript{44,45,47} Extensive research has been dedicated to the structures formed by sequences containing four human telomeric repeats, as this is considered to be the minimum length required for intramolecular G-quadruplex folding. Several G-quadruplex folding topologies have been documented with high resolution structures reported for five intramolecular G-quadruplexes.\textsuperscript{32,33,40,41,48,49}

Crystal and solution structures of various HTelo sequences have been determined and reveal dramatically different topologies (Figure \textit{5 - 2}). Using the 22nt telomeric sequence d[AGGG(TTAGGG)\textsubscript{3}], it was first shown to form an anti-parallel basket-type G-quadruplex in
The same 22 nt sequence, in a K⁺-containing crystal, was found to form a propeller-type parallel-stranded G-quadruplex. More recently, hybrid-type intramolecular G-quadruplex structures have been proposed to be the major conformations in K⁺ solution. Interestingly, these hybrid structures remained even in the co-presentation of high Na⁺. The two hybrid-type telomeric G-quadruplexes are distinct yet related, and appear to be influenced by the 3′-flanking sequence and thus the different capping structures. These are predominantly intramolecular G-quadruplexes termed hybrid-1 and hybrid-2, respectively. Their structures have recently been solved and reveal an identical three-G-tetrad (3 + 1) core structure with differences in the connecting loops. FRET experiments have shown that the two conformations with different half-lives can co-exist and interconvert in solution. The equilibrium between the hybrids was found to be highly dependent on the 3′ sequence of the oligonucleotide and it can be expected to be affected by temperature, ionic conditions and the presence of particular proteins.

Figure 5 - 2: Various folding topologies of the human telomeric 22-mer sequence (HTel22). Reproduced from reference with permission. Grey rectangles represent dG’s in the anti conformation, red rectangles represent dG’s in the syn-orientation. Blue spheres represent dA yellow spheres represent dT.
Since potassium levels in mammalian cells are approximately 150 mM, and generally higher than sodium levels, it is postulated that the K\(^+\) structures are more physiologically relevant. Arguments regarding which structures are more important exist on both sides, some saying that parallel conformations seen in the crystal structure are not biologically relevant and may simply represent an artefact of the crowding conditions induced by the crystalline state, while other authors suggest that molecular crowding conditions, simulated with the introduction of 40% polyethylene glycol, may in fact more accurately represent the *in vivo* situation. Justifications for intramolecular G-quadruplexes predominating over intermolecular forms *in vivo* stem from findings that long telomeric tracts preferentially form into strings of stacked intramolecular G-quadruplexes *in vitro*. However, one cannot discount the potential for the formation of intermolecular G-quadruplexes in some *in vivo* settings, especially if they are modulated through specific protein interactions.

Little is known about the mechanistic details of interconversion of different telomeric G-quadruplex conformations. While the energy difference between the two hybrid-type forms appears to be low, as both conformations coexist in K\(^+\) solution, the kinetics of the interconversion between the two hybrid forms is fairly slow. Furthermore, very few exchange peaks were observed on the NMR time scale, indicating a high energy barrier of the intermediate state(s).

Zhang and coworkers have carried out most of the investigations on the interconversion and have proposed a model for the exchange between the hybrid type K\(^+\) form and the basket type Na\(^+\) form telomeric G-quadruplex structures, through a strand-reorientation mechanism. Addition of K\(^+\) to the preformed Na\(^+\) basket type G-quadruplex readily converts the conformation to the K\(^+\)-form; thus, the two interconvertable K\(^+\) G-quadruplex conformations are both more stable than the Na\(^+\)-basket-type G-quadruplex. An important intermediate in this interconversion was a basket-type intramolecular G-quadruplex with only two G-tetrads but multiple-layer
capping structures formed by loop residues. This two-tetrad basket type structure appears to also fit the transition state of the interconversion of the two hybrid-type conformations. The mechanism of the reorientation, as reported by Zhang, is depicted in Figure 5 - 3.

![Diagram](image)

Figure 5 - 3: The reorganization of HTelo22 from the anti-parallel, basket type structure in Na\(^+\) to the various hybrid forms in K\(^+\). Reproduced from reference with permission.

The 5′-G-strand of the basket-type G-quadruplex (Figure 5 - 3 A) may dissociate from the core G-tetrads (Figure 5 - 3 C) and swing back to the other side of the second G-strand to form a parallel-stranded structural motif with a double-chain-reversal loop (Figure 5 - 3 D). The two-tetrad intermediate appears to fit between the transition from A to C: when the 5′ strand is sliding up to dissociate from the core G-tetrads, a two G-tetrad conformation can be formed transiently. The glycosidic conformations of the tetrad guanines in the two G-tetrad form are the same as those in the hybrid forms (Figure 5 - 3 D/D2). A similar sliding of the 3′-strand (Figure 5 - 3 B2) and dissociation (Figure C2) process is also possible, which results in conversion to the hybrid-2
form (Figure 5 - 3 D2). The two-G-tetrad form is also an intermediate state of the interconversion between the two hybrid type forms. This is supported by the findings that the unfolding processes of the telomeric G-quadruplexes have multiple intermediate states.\textsuperscript{56,57} From these observations, Zhang inferred that the interconversion of the two hybrid forms contained multiple intermediate states, including the two-G-tetrad forms (D $\leftrightarrow$ C $\leftrightarrow$ B $\leftrightarrow$ B2 $\leftrightarrow$ C2 $\leftrightarrow$ D2).

The structural polymorphism becomes increasingly complex within longer HTelo sequences, or when one considers the various proteins associated with the complex. Depending on the DNA context (dsDNA region or ssDNA region), different quadruplex topologies may provide optimal arrangements of higher order telomeric structure. As can be seen in Figure 5 - 4, the hybrid forms seem to permit the compaction of the single stranded region of the telomere,\textsuperscript{33} while the dsDNA region could favour the anti-parallel structures opposite an i-motif.\textsuperscript{58} The parallel topology has also been modeled to provide compaction of the ssDNA component of telomeric DNA.\textsuperscript{40} The simpler fold of parallel structures, compared with anti-parallel/hybrid structural models, suggests an obvious pathway for readily folding and unfolding G-quadruplex structures. The loops of successively stacked parallel quadruplexes result in exterior loops of the superhelix, suitable for interaction with telomeric proteins or with the nuclear envelope, or for inhibiting telomerase extension, especially when stabilized by ligand binding.\textsuperscript{40}
Figure 5 - 4: (a) Possible model of duplex, i-motif and G-quadruplex association for the HTelo repeat, reproduced from reference with permission. (b) The hybrid-type telomeric G-quadruplex structures can be readily folded and stacked end to end to form compact-stacking structures for multimers in the elongated telomeric DNA, reproduced from reference with permission. (c) Parallel-type telomeric G-quadruplex structures folding and stacked end to end with same direction or alternate direction stacking for connection between G-quadruplex blocks, reproduced from reference with permission.

5.2.3. Telomerase Therapeutics

Due to the fact that most normal human cells are telomerase-silent, while telomerase activity is detected in nearly all cancers, telomerase has emerged as an almost universal target for cancer therapeutics. Compared to conventional chemotherapeutic approaches, telomerase-based therapies could possess greater specificity, lower toxicity and reduced side effects. There are three general classes of agents that have been developed to target telomerase biology; gene
therapy, immunotherapy and small molecule inhibitors. For the purpose of this chapter the focus will be on small molecule inhibitors.

5.2.3.1. G-Quadruplex Stabilizing Molecules

It was first shown in 1991 by Zahler et. al that an intramolecular telomeric G-quadruplex could not be extended by Oxytricha telomerase in vitro.\textsuperscript{19} Since telomerase is active in most human cancers, and its activity is influenced by G-quadruplex structures, a variety of small molecule ligands with different specificities and target regions that bind and stabilize G-quadruplex structures are being tested in various assays.\textsuperscript{39,60,61} The goal is to find ligands that promote the formation of certain types of telomeric quadruplexes in vivo which inhibit telomerase by preventing annealing of telomerase RNA to G-strand overhangs.\textsuperscript{19,20} A tremendous amount of effort has been devoted to in vitro studies for ligand-induced stabilization of intramolecular telomeric G-quadruplexes.\textsuperscript{19,39,62,63} One prominent challenge in this area is the structural elucidation of folding topologies adopted by the oligonucleotides upon ligand-induced stabilization of the quadruplex. Ligands that have been demonstrated to inhibit telomerase in a direct assay with high specificity for G-quadruplex over duplex DNA (30-100 fold) include: BRACO19,\textsuperscript{64} 360A and 307A,\textsuperscript{65} Phen-DC\textsuperscript{66} and telomestatin (Figure 5 - 5).\textsuperscript{67} NMM has been shown to be a quadruplex-stabilizing ligand that is specific for a parallel topology.\textsuperscript{52}
Since there are a large number of non-telomeric, G-quadruplex-forming sequences in the human genome, an important consideration when evaluating potential telomere-targeting drugs is their specificity for particular G-quadruplex conformations. Some porphyrin ligands have been found to interact with telomeric G-quadruplexes, with minimal degree’s of specificity over interaction with a G-quadruplex in the promoter of the c-Myc oncogene. Furthermore, G-quadruplex formation at the RNA component of human telomerase was shown to play a role in telomerase assembly, which actually stimulated telomere maintenance in vivo. This shows that ligands with broad quadruplex recognition, originally designed to inhibit telomerase, could increase telomerase processivity and genome instability that could lead to secondary tumors.

While G-quadruplex stabilising molecules are showing great promise as anti-cancer drugs, their mechanisms of cellular action and the likelihood of adverse effects on healthy, proliferating cells requires further investigation, which is an avenue to pursue for academic research. The mechanism of action for many G-quadruplex ligands resulting in telomerase
inhibition has been reviewed.\textsuperscript{71} It was reported that most ligands primarily act by locking telomerase substrates into a folded G-quadruplex conformation. However, it was also noted that telomerase inhibition was achieved by mechanisms that did not involve G-quadruplex formation, such as interactions with single-stranded primers or telomerase core components.\textsuperscript{71} Inhibition of telomerase activity in control experiments reinforced the notion that quadruplex ligands act as broad inhibitors of telomere-related processes. Although these molecules preferentially localize and act at telomeres, they have been found to inhibit the activity of canonical DNA polymerases\textsuperscript{72} and helicases\textsuperscript{73} on telomeric substrates, and delocalize telomeric proteins from telomeres,\textsuperscript{74} which contributes to their cytotoxicity toward cancer cells and their efficiency as anticancer drugs.

While attention has been paid to inducing formation of G-quadruplexes within HTelo sequences, there exists another scenario that less focus has been paid to: what is preventing G-quadruplex formation or unfolding the structures once formed? Thermodynamic and kinetic stabilities of G-quadruplex structures suggest that their formation \textit{in vivo} should be favoured;\textsuperscript{36,75} therefore, active mechanisms must exist to prevent their formation, or to unfold them once formed. Studies by Zaug and Paeschke have found that certain TEBP’s disrupt G-quadruplexes, forming stable complexes that permit telomerase extension.\textsuperscript{27,36} Additional work by Huber et al. has shown G-quadruplex unwinding by specific helicases.\textsuperscript{76} Interestingly, this unwinding was inhibited in the presence of N-methyl-mesoporphyrin (NMM); however, it is unclear which quadruplex structures were formed initially and which ones can be unwound. Work by Nicolaudis has found that NMM has the unique ability to recognize parallel-stranded but not anti-parallel G-quadruplex structures.\textsuperscript{52} This infers a parallel topology was resolved by the helicases, since NMM only interacts with parallel quadruplexes, it likely bound to the end of the parallel quadruplex and blocked the activity of the helicases. Further studies are required to determine if this is specific for parallel topologies or if the hybrid forms or anti-parallel topologies can also be resolved.
FurdG probes within HTelo oligonucleotides provide a tool that can be exploited to answer these questions. The placement of FurdG probes within the G-tetrad of the quadruplex provides direct probing capabilities for G-quadruplex unwinding, as evidenced within the TBA oligonucleotides previously discussed. This chapter aims to determine whether FurdG probes are able to distinguish between the various folding topologies of the G-quadruplexes formed by the HTelo sequence. The conformational heterogeneity of the HTelo sequence will allow further characterization of the specific impacts imparted via modifications of a G-tetrad. The FurdG probe has been shown to provide fluorescent diagnostic signals to distinguish between various DNA tertiary structures (i.e. single strand vs. duplex vs. quadruplex). If FurdG can be shown to distinguish various G-quadruplex structures it may become useful in the search for small molecules and proteins that preferentially stabilize specific G-quadruplexes. Furthermore, incorporation of FurdG within HTelo will provide insight into the effects of C8-dG adduct formation within the telomeric region of the chromosome; it is expected that DNA damage in this region impacts the complex equilibria between duplex/quadruplex, single-strand/quadruplex, or different quadruplex structures.

5.3. Materials and Methods

5.3.1. General Methods

A detailed description of the general methods for experiments pertaining to this chapter can be found in Appendix A.

5.3.2. Oligonucleotide Synthesis and Purification

Preparation of the fully protected FurdG phosphoramidite was performed as outlined previously in Scheme 2 - 5. To study the impact of the FurdG adduct on quadruplex structure and determine the probe’s performance within different positions of the Human Telomeric (HTelo) sequence, it was incorporated into the 22-mer HTelo sequence (5'-AGGG(TTAGGG)₃) using
solid phase oligonucleotide synthesis. The positions chosen for modification included G3, G8 and G10 in order to place Fur-dG in each tetrad of the corresponding G-quadruplex.

5.3.3. Mass Spectrometry Analysis

Following HPLC purification, fractions containing only pure material was pooled and lyophilized to dryness and analyzed by electrospray ionization mass spectrometry (ESI-MS) as described in Appendix A. The ESI spectra showed the expected clusters of multiply charged peaks for the modified oligonucleotides, with results of this analysis summarized in Table 5 - 1. Mass spectra for all C8-aryl-dG modified oligonucleotides are included in Appendix C.

Table 5 - 1: ESI-MS analysis of mHTelo22 oligonucleotides

<table>
<thead>
<tr>
<th>mHTelo</th>
<th>Product Formula</th>
<th>Calcd mass</th>
<th>Exptl m/z</th>
<th>Exptl Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fur-dG</td>
<td>C_{224}H_{373}N_{92}O_{133}P_{21}</td>
<td>7031.2</td>
<td>[M-8H]^{8-} = 877.9</td>
<td>7031.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-9H]^{9-} = 780.3</td>
<td>7031.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-10H]^{10-} = 702.1</td>
<td>7031.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[M-11H]^{11-} = 638.5</td>
<td>7031.2</td>
</tr>
</tbody>
</table>

Upon confirmation via mass spectrometry, sample concentrations were quantified according to procedures outlined in Appendix A.

5.4. Results and Discussion

5.4.1. Impact of 8-aryl-dG Modifications on Structure and Stability

Depicted in Scheme 5 - 2 are the various quadruplexes that HTelo22 may form, outlining the position of modification within each topology.
5.4.1.1. Thermal Melting Analysis

Thermal Melting experiments were conducted in order to assess the positional impact that \( \text{Fur} \text{dG} \) modification had on the overall stability of the mHTelo oligonucleotides. Results are summarized in Table 5 - 2 and depicted in Figure 5 - 6.

Table 5 - 2: Tabulated \( T_m \) values for \( \text{Fur} \text{dG} \) mHTelo oligonucleotides folding into duplex or quadruplex structures.

<table>
<thead>
<tr>
<th>Nucleoside</th>
<th>Position</th>
<th>Position</th>
<th>( \Delta T_{m} )</th>
<th>( \Delta T_{m} )</th>
<th>( \Delta T_{m} )</th>
<th>( \Delta T_{m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Na(^+)</td>
<td>Duplex</td>
<td>Na(^+)</td>
<td>Duplex</td>
<td>Na(^+)</td>
</tr>
<tr>
<td>dG</td>
<td>/</td>
<td>69.7 ± 1.7</td>
<td>/</td>
<td>61.6 ± 1.2</td>
<td>/</td>
<td>68.0 ± 0.9</td>
</tr>
<tr>
<td>( \text{Fur} \text{dG} )</td>
<td>3</td>
<td>65.3 ± 1.4</td>
<td>-4.4</td>
<td>68.3 ± 1.1</td>
<td>6.7</td>
<td>71.3 ± 1.6</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>64.7 ± 1.7</td>
<td>-5</td>
<td>69.4 ± 1.2</td>
<td>7.8</td>
<td>80 ± 3.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>65.3 ± 1.7</td>
<td>-4.4</td>
<td>64.4 ± 1.1</td>
<td>2.8</td>
<td>68.5 ± 1.0</td>
</tr>
</tbody>
</table>
The utilization of \( \text{Fur} \text{dG} \) within the HTelo sequence was based on several realizations from the mTBA studies. Firstly, \( \text{Fur} \text{dG} \) is able to adopt a \textit{syn} or \textit{anti} orientation within G-quadruplex structures. Quadruplex folding studies within the mTBA scaffold showed that \( \text{Fur} \text{dG} \) stabilized quadruplexes when \textit{syn}-G’s were modified, while modification at \textit{anti}-G’s destabilized (but did not inhibit) quadruplex folding. This is an important attribute, as the structural polymorphism demonstrated within the HTelo sequence requires certain G’s to switch from \textit{syn} to \textit{anti} (and vice versa).

The \( T_m \) data presented in Table 5 - 2 demonstrates that \( \text{Fur} \text{dG} \) modifications destabilized duplex folding by approximately 5°C. This is in agreement with the previously reported m\textit{NarI} and mTBA oligonucleotides; the \textit{syn} preference of the \( \text{Fur} \text{dG} \) nucleoside destabilizes dsDNA as the probe is forced to attain an \textit{anti} conformation in order to participate in W-C H-bonding with its complementary dC base within the duplex.
Quadruplex stabilization is clearly positionally and situationally distinct. With regard to the Na\(^+\)-stabilized quadruplexes, it is reported that the basket-type intramolecular G-quadruplex is preferred which has loop-diagonal-loop orientations of the strands (Scheme 5 - 2).\(^{57}\) In this structure G3, G8 and G10 are all syn, and changes in \(T_m\) are independent of the glycosidic orientation. Indeed, FurdG modifications at G3 and G8 are within experimental error (\(\Delta T_m\)'s 7-8°C); however, G10 did not exhibit the same level of stabilization (\(\Delta T_m = 2.8^\circ\)C). The major difference between the positions in this sequence concerns G10, which is involved in the edgewise diagonal loop, while G3 and G8 are not (G3 is the middle tetrad and G8 is the edgewise loop). There may be some unfavourable steric interactions imposed by an edgewise loop, as the TTA-loop structure needs to diagonally cross the tetrad core, which is a longer distance than in an edgewise loop, where it loops across one side of the tetrad. Furthermore, it seems plausible that W-C H-bonding interactions can form more readily within an edgewise loop, where the terminal T and A of the TTA loop are close enough to form H-bonds. Base stacking interactions may also contribute to the lack of stability at G10, which is stacked by one purine (dG) and one pyrimidine (T) (5'-GG\(_{10}\)T), opposed to G3 (5'-GG\(_3\)G), and G8 (5'-AG\(_8\)G), which are stacked by two purines. As previously described, purine-purine stacking interactions offer the greatest potential for \(\pi\)-overlap of the aromatic systems, and typically result in larger stacking interactions.\(^{78}\) It is clear that MD simulations would provide valuable insight into the structural analysis of quadruplex structures within HTelo oligonucleotides.

The analysis of the K\(^+\)-stabilized quadruplexes is more complex. Furthermore, the \(T_m\) profiles generated in K\(^+\) were not as uniform as the ones generated in Na\(^+\) and exhibited hysteresis. This is typically encountered when a two-state equilibrium between folded and unfolded oligonucleotides is encountered, or when there are multiple intermediates involved in the transition.\(^{79}\) The K\(^+\)-stabilized HTelo sequence is highly polymorphic and multiple folding topologies exist in equilibrium.\(^{43}\) Of these forms, the most prevalent in solution are the anti-
parallel basket-type G-quadruplex, which is the same as the Na\(^{+}\)-stabilized structure, along with two hybrid forms (denoted hybrid forms 1 and 2), which are mixed parallel-anti-parallel G-quadruplexes (Scheme 5 - 2).\(^{55}\) The hybrid forms have been reported to be the dominant structures prevalent in K\(^{+}\) solution.\(^{45}\) The unmodified HTelo sequence displayed a 6.4°C stabilization upon replacement of Na\(^{+}\) with K\(^{+}\); this is in agreement with published results indicating that the K\(^{+}\) hybrid form topologies are favoured in solutions containing both Na\(^{+}\) and K\(^{+}\).\(^{45}\) We expected this equilibrium to be affected upon incorporation of Fur\(\text{dG}\) at G3, G8 or G10. Indeed, the position of modification resulted in specific changes in K\(^{+}\)-stabilized quadruplexes that impacted the equilibrium of topologies. \(T_m\) analysis depicted universal stabilization (Table 5 - 2), with stabilization increasing in the order: \(\text{Fur\(\text{dG}\)@10} (\Delta T_m = 0.5°C) < \text{Fur\(\text{dG}\)@3} (\Delta T_m = 3.3°C) << \text{Fur\(\text{dG}\)@8} (\Delta T_m = 12°C).

The orientation of the glycosidic bond changes at G10 and G3 when the basket type quadruplex (\textit{syn} orientation) rearranges to one of the hybrid forms (\textit{anti} orientation) (Scheme 5 - 2). One would expect incorporation of a Fur\(\text{dG}\) modification at these positions to result in destabilization of the hybrid forms (in accordance with our Fur\(\text{dG}\) studies within mTBA one would expect a \(\Delta T_m\) of approximately -11°C).\(^{77}\) This is evidenced by the difference in \(T_m\) values when comparing the K\(^{+}\) and Na\(^{+}\)-stabilized quadruplexes, as only marginal stabilities for \(\text{Fur\(\text{dG}\)@3} (\Delta T_m \text{ K}^{+}\text{-Na}^{+} = 3.3) \text{ and Fur\(\text{dG}\)@10} (\Delta T_m \text{ K}^{+}\text{-Na}^{+} = 4.1) \) were observed. In this case, the unmodified HTelo sequence is stabilized by 6.4°C, because the hybrid forms are preferentially formed in K\(^{+}\). For the \(\text{Fur\(\text{dG}\)@3} \) and \(\text{Fur\(\text{dG}\)@10} \) mHTelo oligonucleotides, attainment of the hybrid topology is not as favourable, as the orientation of the glycosidic bond is required to change from \textit{syn} to \textit{anti}, and equilibria between hybrid forms and the anti-parallel basket form exists. It is possible that the amount of quadruplexes folding into an anti-parallel basket type remains greater than hybrid forms in K\(^{+}\), owing to the Fur\(\text{dG}\) probes preference for a \textit{syn}-glycosidic orientation.
Fur\textsubscript{dG}\textsubscript{@8} exhibited much larger stabilization for the $T_m$ analysis in K$^+$ ($\Delta T_m = 12^\circ$C). The $T_m$ of this solution was so high that the fully dissociated quadruplex was not observed in the $T_m$ studies; increasing the temperature to 95$^\circ$C was not sufficient to generate a fully dissociated baseline. Attempts to heat the sample further were unsuccessful. The transition was clearly evident, and the $T_m$ value is reliable within 3$^\circ$C error. Position 8 remains in a syn-orientation throughout the transition from the basket type G-quadruplex to either of the hybrid forms (Scheme 5 - 2). Therefore, it would be expected that the Fur\textsubscript{dG}\textsubscript{@8} mHTelo oligonucleotides are able to attain the hybrid topology more readily than the Fur\textsubscript{dG}\textsubscript{@3} or Fur\textsubscript{dG}\textsubscript{@10} derivatives. This is supported by the $\Delta T_m$ comparing the K$^+$ and Na$^+$ quadruplexes ($\Delta T_m$ K$^+$-Na$^+$ = 10$^\circ$C), which is larger than the Fur\textsubscript{dG}\textsubscript{@3} ($\Delta T_m$ K$^+$-Na$^+$ = 3$^\circ$C) or Fur\textsubscript{dG}\textsubscript{@10} ($\Delta T_m$ K$^+$-Na$^+$ = 4.1$^\circ$C) derivatives. It is likely that the anti-parallel topology is still present in solution, due to the large degree of stabilization exhibited for the anti-parallel basket type quadruplex in Na$^+$ ($\Delta T_m$ = 7.8$^\circ$C); however, there is clearly an energy benefit for attaining the hybrid topology. Due to the fact that the glycosidic bond orientation remains in a syn-orientation for G8, the amounts of each topology in equilibrium would be dependent on steric and base stacking interactions within the different loop structures.

Modification of HTelo with Fur\textsubscript{dG} clearly results in duplex destabilization and quadruplex stabilization. The $\Delta \Delta T_m$ ($\Delta T_m$ K$^+$ quadruplex - $\Delta T_m$ Na$^+$ duplex) comparing quadruplex stabilization over duplex upon incorporation of Fur\textsubscript{dG} is between 5 and 15$^\circ$C for a single Fur\textsubscript{dG} modification depending on position. This has toxicological implications, as one could expect damage within double stranded telomeric DNA to result in the region preferring to form quadruplex structures, especially if numerous adduct’s were formed in close proximity. DNA repair is typically inhibited within the telomeric region, which prevents chromosomal fusions and genome instability.\textsuperscript{80} It has been found that double strand breaks next to telomeric repeats resist DNA repair.\textsuperscript{81} Telomeres also have been shown to accumulate DNA damage stemming from
alkylation, oxidative and radical stress, as well as UV exposure. Rochette and coworkers studied the accumulation of cyclobutane pyrimidine dimers from UV light exposure, they reported that telomeres were unique in terms of their biophysical UV sensitivity, their prevention of excision repair, and their tolerance of unrepaired lesions. Cells containing persistently high levels of telomeric cyclobutane pyrimidine dimers eventually proliferated; however, chronic UV irradiation of cells did not accelerate telomere shortening. The authors suggested that by utilizing a lesion-tolerance strategy, rather than a repair strategy, double strand breaks at closely-opposed excision repair sites would be prevented.

Within the single stranded portion of the telomeric region, adduct formation will significantly alter the quadruplexes attainable folding topologies. The toxicological implications of this are completely unknown, the roles of the various topologies have not been uncovered. Utilization of Fur'dG modified HTelo oligonucleotides would greatly impact this field and provide researchers the tools needed to determine how various quadruplex structures favour, or inhibit, certain biological processes. Further characterization of the impact of Fur'dG modification on HTelo will be presented, aimed at determining the impact on quadruplex structure.

5.4.1.2. Circular Dichroism

CD was carried out on all modified duplexes and quadruplexes to determine how Fur'dG impacted oligonucleotide folding into corresponding duplexes and quadruplexes. Spectra were monitored at 6 μM concentrations and data are presented in Figure 5 - 7.

CD interpretation of the various HTelo G-quadruplexes can become quite convoluted. Recall that for CD analysis of G-quadruplexes interpretation can be summarized in the following way: (1) a band at approximately 290 nm is indicative of the presence of stacking of guanosines of different glycosidic bond angles (GBA); (2) a positive band at approximately 240 nm coupled with a negative band at approximately 260 nm indicates the absence of stacking bases of the same
GBA in the quadruplex; (3) a negative band at approximately 240 nm coupled with a positive band at approximately 260 nm indicates the presence of stacking of guanosines of the same GBA

Figure 5-7: Circular Dichroism spectra for mHTelo oligonucleotides. Black line depicts unmodified HTelo oligonucleotide while the red line depicts Fur$^{\text{dG@3}}$, blue line depicts Fur$^{\text{dG@8}}$, and green line depicts Fur$^{\text{dG@10}}$. (a) Na$^+$ stabilized duplex; (b) Na$^+$ stabilized quadruplex; (c) K$^+$ stabilized quadruplex
All duplex CD spectra recorded (Figure 5 - 7 (a)) displayed the characteristic positive long wavelength band between 260 and 280 nm, and the negative band around 245 nm attributed to B-form DNA (i.e. stacking of guanosines with same GBA).\textsuperscript{84} For no obvious reason, the amplitude at approximately 220 nm was highly variable, as the unmodified and Fur\textsubscript{d}G@G8 sequences displayed a negative peak while Fur\textsubscript{d}G@G3 and Fur\textsubscript{d}G@G10 each had a positive peak.

Spectra of the Na\textsuperscript{+}-stabilized quadruplexes (Figure 5 - 7 (b)) contain typical signals for an anti-parallel-G quadruplex. These display a positive band at 290 nm (stacking of guanosines of different GBA), as well as a positive band at 240 nm coupled with a negative band at 260 nm, which indicates the absence of stacking bases of the same GBA in the quadruplex.\textsuperscript{85} All modifications within HTelo resulted in similar CD spectra, with overall amplitudes slightly smaller than the unmodified sequence. This slight reduction is most likely a result of reduced base stacking interactions due to small local perturbations imparted by the Fur\textsubscript{d}G modification.

The CD spectra of K\textsuperscript{+}-stabilized quadruplexes were more variable, as expected due to the conformational heterogeneity associated with K\textsuperscript{+}-stabilized G-quadruplexes previously reported.\textsuperscript{32, 53, 78} In K\textsuperscript{+} solution, the spectrum of the unmodified sequence showed a prevalent positive peak at 290 nm (stacking of G’s with different GBA), with a shoulder at 270 nm (hybrid form) and slight positive amplitudes down to about 248 nm. Below this point, they turned negative, which shows the presence of stacking of G’s of the same GBA. This spectrum was similar to those reported from the human telomere DNA in K\textsuperscript{+} solution, which was attributed to an equilibrium of conformations comprised of hybrid and anti-parallel topologies, for which the hybrid forms were more stable and more prevalent.\textsuperscript{32,54,86}

Modification at G3 (Figure 5 - 7 (c) red trace) resulted in a spectrum that most closely related to an anti-parallel G-quadruplex, as the shoulder at 270 nm was replaced with a negative peak at approximately 260 nm and the 240 nm peak grew in intensity. As previously mentioned,
it is plausible that modification at G3 would shift the conformational equilibrium of G-quadruplexes in K⁺ solution from either hybrid form to an anti-parallel form, as G3 is syn in the anti-parallel conformation and anti in the hybrids.

Modification at G8 resulted in slight changes from the unmodified sequence: the spectrum displayed larger amplitudes at 290 nm and 270 nm, which remained positive down to approximately 240 nm with a small trough at approximately 260 nm. Here, it seems that the modification has increased the abundance of hybrid topologies, as the CD suggests both stacking interactions of the same GBA and different GBA have increased.

Modification at G10 was expected to display a similar spectrum to G3; however, it seems as though Fur⁰dG@10 modified oligonucleotides still adopt a mixture of anti-parallel and hybrid topologies in solution. It is evident that Fur⁰dG@10 yields a larger positive amplitude at 290 nm and has the 240 nm peak coupled with a deeper trough at 260 nm (which suggests more anti-parallel conformation); however, the 260 nm trough remains positive overall, possibly due to the presence of the 270 nm shoulder which has been attributed to the hybrid topology. This correlates with the \( T_m \) analysis, if one considers that the \( \Delta T_m \) K⁺ - Na⁺ value for Fur⁰dG@10 (+4.1°C) is larger than the Fur⁰dG@3 (+3°C), while the \( \Delta T_m \) of Fur⁰dG@10 in Na⁺ was the smallest among all modifications (+2.8°C). It was previously suggested that Fur⁰dG@10 had unfavourable interactions with the diagonal loop structure in the anti-parallel basket. Transformation to a hybrid form eliminates this interaction (forms edgewise loops), and allows the hybrid topology to compete with the formation of the anti-parallel topology, even though the glycosidic bond orientation is changing from syn to anti.

Another possible reason G10 can attain an anti conformation while G3 cannot stems from the fact that G3 is in the middle tetrad, while G10 is in the terminal tetrad. This results in less stacking interactions for Fur⁰dG@10, and could allow for the Fur⁰dG to attain an anti conformation.
more readily. This is also supported by the CD spectrum observed for $\text{Fur}dG@ G8$, which more closely resembled the modification at G10 than G3, although it is clear that G8 favoured the hybrid forms more than G10 and G3. Since G8 remains in a syn orientation for the hybrid topologies, it preferentially adopts the hybrid forms that are stabilized in $K^+$. In summary, it is evident that the equilibrium of topologies HTelo folds into is impacted by $\text{Fur}dG$ modification. $\text{Fur}dG@8$ is in equilibrium between hybrid and anti-parallel topologies with a preference for the hybrid forms. $\text{Fur}dG@10$ also exists in an equilibrium between hybrid and anti-parallel topologies; however, it seems to have slightly more preference for the anti-parallel topology. $\text{Fur}dG@3$ has a strong preference for the anti-parallel topology.

5.4.2. Fluorescence Response

The HTelo22 sequence is conformationally diverse, and crystallography and NMR have produced divergent high-resolution structures. In order to distinguish between the various conformations under conditions similar to those that prevail in vivo, low-resolution methods, in conjunction with fluorescent nucleotide modifications, have been employed. Fluorescent probes can generally serve as versatile, site-specific conformational reporters. The use of site-specific spectroscopic probes is particularly useful because of the ability to report specific, rather than global, structural differences brought about changes in solution conditions.

Some published studies have used 2-AP, which was substituted into various loop positions within the HTelo22 sequence. Kimura et al. have substituted 2-AP for A7, A13 and A19 in the HTelo22 sequence. They report that modification at position A7 resulted in a significant increase in fluorescence intensity upon a duplex-quadruplex transformation. More recently, Gray and coworkers have used 2-AP fluorescence to assess the dynamics and equilibria for folding of specific loops within HTel22. They report that Htel22 folding induced by Na$^+$ or K$^+$ proceeded through at least three kinetic steps. Furthermore, from the observation of K$^+$-
dependent folding, they inferred the presence of different conformational ensembles in different K\textsuperscript{+} concentrations.\textsuperscript{86}

An alternative use of fluorescent modified nucleosides within HTelo22 is the modification of various positions within the G-tetrads (as opposed to loop regions). Through the use of \textsuperscript{Fur}dG nucleosides, it has clearly been shown that the position of modification has implications regarding the overall equilibrium of topologies in solution. Nevertheless, it has been shown that \textsuperscript{Fur}dG can adopt syn or anti conformations, which permits its use as a probe in this sequence. To determine the probes’ performance within various topologies, the fluorescence response of \textsuperscript{Fur}dG was measured for the \textsuperscript{Fur}dG@3, \textsuperscript{Fur}dG@8 and \textsuperscript{Fur}dG@10 mHTelo oligonucleotides; spectra are depicted in Figure 5 - 8 while data are tabulated in Table 5 - 3.
Figure 5 - 8: Excitation and Emission spectra ($\lambda_{ex} = 320$ nm, $\lambda_{em} 380$ nm) of 6 μM solutions $^{\text{Fur}}dG$ mHTelo22 within (a) duplex, (b) Na$^+$ stabilized quadruplexes, (c) K$^+$ stabilized quadruplexes. Red line corresponds to $^{\text{Fur}}dG@3$; blue line corresponds to $^{\text{Fur}}dG@8$; green line corresponds to $^{\text{Fur}}dG@10$. 
Table 5 - 3: Tabulated photophysical properties of mHTelo oligonucleotides.

<table>
<thead>
<tr>
<th>Position</th>
<th>Scenario</th>
<th>$\lambda_{\text{max}}$ (nm)</th>
<th>$\lambda_{\text{em}}$ (nm)</th>
<th>Emission Intensity</th>
<th>$I_{\text{rel}}$ (Q/D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Na$^+$ Quadruplex</td>
<td>249.1 / / /</td>
<td>290.0 378.9 928.6 14.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quadruplex</td>
<td>250.0 / / /</td>
<td>287.0 380.0 597.9 9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>315.1 386.9 63.8 / /</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Na$^+$ Quadruplex</td>
<td>249.1 / / /</td>
<td>289.1 378.9 959.8 35.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quadruplex</td>
<td>250.0 / / /</td>
<td>289.1 380.9 362.8 13.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>318.0 381.8 250.7 9.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Na$^+$ Quadruplex</td>
<td>250.0 / / /</td>
<td>292.0 378.9 370.7 39.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>K$^+$ Quadruplex</td>
<td>250.0 / / /</td>
<td>289.1 388.9 117.3 12.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Duplex</td>
<td>312.0 384.9 9.5 / /</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall, the modified duplexes displayed quenched fluorescence emission. This is consistent with the Fur$dG$ probe residing in a duplex environment, as evidenced for the mNarI and mTBA oligonucleotides previously discussed. For the mHTelo22 duplexes, the order of fluorescence intensity is $\text{Fur}dG@10 < \text{Fur}dG@8 < \text{Fur}dG@3$. $\text{Fur}dG$ is expected to primarily attain an anti-conformation within these duplexes (which maintains W-C H-bonding), although a syn orientation cannot be ruled out; as previously mentioned for the mNarI sequences, $\text{Fur}dG$ modifications were able to attain some syn orientation within the duplex, as MD simulations predicted $\text{Fur}dG$ to have the lowest barrier to rotation about its glycosidic bond (14 kJ/mol) opposite C. Base sequence context (Figure 5 - 3) for HTelo22 results in $\text{Fur}dG@3$ being flanked by two dG residues; while $\text{Fur}dG@8$ is flanked by one dG and one dA, and $\text{Fur}dG@10$ is flanked
by one dG and one dT. Base stacking interactions are able to modulate the conformational preference of Fur-dG. Recall that there is some degree of twist between the two aryl systems for 8-aryl-dG nucleosides (sterics prevent coplanarity); therefore, due to the small size of the furanyl moiety, it is the dG component of the adduct that has the larger π-surface and would prefer to stack with neighbouring purines. This results in further preference of the anti conformation, for which the furanyl moiety is localized to the major groove. If this is the case, it seems as if increased stacking interactions between Fur-dG and the flanking purines have resulted in larger fluorescence emission intensities. This contrasts with the fluorescence response produced by CNPh-dG, for which increased base stacking interactions leads to quenching of the fluorescence (by PET with dG). It would be informative to incorporate CNPh-dG nucleosides within HTelo22 to probe such scenarios with an electronically distinct C8-aryl-dG derivative. The emission maximum also supports Fur-dG being in an anti conformation and exposed to the aqueous environment, as the duplexes emission is relatively red shifted compared to the nucleoside; this was also observed with the NarI and mTBA dsDNA structures. Specifically, changes in emission wavelength may offer some insight into the equilibrium of anti/syn orientations as the most red shifted emission occurs for Fur-dG@3 followed by Fur-dG@8 and Fur-dG@10.

Fur-dG modified quadruplexes have distinct excitation spectra, containing a diagnostic peak at 290 nm indicative of anti-parallel quadruplex folding. In the Na+-stabilized quadruplexes, emission intensity was positionally dependent resulting in a trend of Fur-dG@10 < Fur-dG@3 < Fur-dG@8. All of these quadruplexes were shown via CD to attain the anti-parallel basket type structure in solution (Figure 5 - 2). Recall that Fur-dG@10 exhibited the lowest degree of stabilization in T_m analysis (Table 5 - 2) while Fur-dG@3 was similar to Fur-dG@8. As depicted in Figure 5 - 8 (b), Fur-dG@10 was the least emissive position within the Na+ stabilized quadruplexes. This correlates with the hypothesis that modification at position 10 resulted in the lowest degree of stacking interactions and had higher degrees of steric clash with the loop
structures. Fur\textsuperscript{dG@8} is more emissive than Fur\textsuperscript{dG@3} because it is more exposed to an aqueous environment. Fur\textsuperscript{dG@3} is sandwiched between two G-tetrads, and only the furanyl moiety would be exposed to H\textsubscript{2}O. Fur\textsuperscript{dG@8} is on a terminal tetrad, which results in significant exposure of one face of the adduct to H\textsubscript{2}O, along with the furanyl moiety. Additionally, differences in rigidity experienced by Fur\textsuperscript{dG} modifications can result in changes in fluorescence emission. It has been noted that increased solvent viscosity results in excited states maintaining geometries similar to that of the ground state.\textsuperscript{88} In viscous solvents, molecular rotation is inhibited, which may cause excited states to resemble ground states. With Fur\textsuperscript{dG}, increasing solvent viscosity reduces fluorescence intensity.\textsuperscript{89} Restricted motion should also influence the emissivity of Fur\textsuperscript{dG} in the HTelo sequence, and this effect should vary with position. Fur\textsuperscript{dG@8} would be expected to experience a less rigid local environment in the terminal tetrad than Fur\textsuperscript{dG@3}, which is within the internal G-tetrad. The fluorescence response of the Fur\textsuperscript{dG} nucleoside within the anti-parallel basket-type quadruplex has allowed us to probe the local environment within the quadruplex.

Interpretation of the fluorescent signals generated by the K\textsuperscript{+}-stabilized quadruplexes is not trivial, since HTelo22 exists as a mixture of unknown topologies in unknown proportions.\textsuperscript{46} The spectroscopic signals observed for such solutions will represent an average of the signals for each conformer weighted for abundance. Each conformation would be expected to generate a specific fluorescence emission; however, the signals will be further dispersed by the local micro-heterogeneity associated with the equilibrium of folded structures in the presence of K\textsuperscript{+}.

Some useful information can be derived from the fluorescence response of Fur\textsuperscript{dG} when combined with the CD signals presented in Figure 5 - 7. The Fur\textsuperscript{dG@3} modification favoured the anti-parallel, basket type quadruplex in K\textsuperscript{+} solution. The fluorescence response of Fur\textsuperscript{dG} agrees with the CD data, as emission intensity is very similar between the Na\textsuperscript{+} and K\textsuperscript{+}-stabilized quadruplexes (I\textsubscript{rel} Na\textsuperscript{+} vs. K\textsuperscript{+} is < 1.2); the small differences in emission intensity may be attributed to the counter-ion present as reported in the literature for 2-AP.\textsuperscript{43}
Fur\textsubscript{dG}@8 and Fur\textsubscript{dG}@10 modifications exhibit quenched emission, likely due to the higher proportion of hybrid topologies as evidenced by CD (Figure 5 - 7). Fur\textsubscript{dG}@8 is quenched by a factor of 1.6 when going from the Na\textsuperscript{+} counter ion to K\textsuperscript{+}, while Fur\textsubscript{dG}@10 was quenched by a factor of 2.7. It is difficult to work out the exact mechanism of quenching, since the sample is so diverse structurally. CD spectra in K\textsuperscript{+} solution suggests that Fur\textsubscript{dG}@8 and Fur\textsubscript{dG}@10 quadruplexes exist in larger amounts of hybrid topologies, compared to Fur\textsubscript{dG}@3, this structural change may provide fluorescence quenching.

5.4.3. Modulation of Folding Topologies

Small molecule co-solvents or macromolecular co-solutes are commonly used to perturb conformational equilibria. There are four possible mechanisms by which such additives may influence conformation: macromolecular crowding, which acts on differences in shape and size; osmotic stress, which changes hydration conditions; differential binding, general solvation effects; and by general electrostatics, for example dielectric effects on ionic activity.\textsuperscript{90}

It has been demonstrated that HTelo DNA can adopt a parallel-stranded conformation in physiological K\textsuperscript{+} solution induced under various conditions. Specifically, the parallel conformation has been documented in solutions under dehydrating conditions,\textsuperscript{91,92} in high concentrations of polyethylene glycol (PEG)\textsuperscript{91} and with various co-solutes.\textsuperscript{52,54,93} Much work has been done with crowding agents, the use of which intends to mimic the molecular crowding conditions that are present in biological systems, as opposed to using the dilute conditions present in most \textit{in vitro} studies.

Hydration significantly impacts macromolecule stability and conformation.\textsuperscript{94} The structure of water in the cell is inhomogeneous, as the hydration layers around macromolecules differ from bulk water. Macromolecular folding reactions that involve changes in the extent of hydration can be influenced by water, such that a reduction in water activity favors the less
hydrated species. Buscaglia and coworkers report that the 22nt HTelo quadruplex will undergo a transition from a hybrid conformation to a parallel propeller conformation upon addition of acetonitrile or ethanol. They proposed that this conformational change was driven by decreased water activity, which would imply that the propeller form is less hydrated than the hybrid form.

5.4.3.1. Response to Water Activity

The effect of water activity was tested on the Fur-dG mHTelo oligonucleotides. The site-specific impact of Fur-dG modifications on the conformational equilibrium was probed via CD, while the ability of Fur-dG to distinguish the anti-parallel/hybrid conformations from parallel propeller conformations was determined via fluorescence spectroscopy. This was tested by making 6 μM quadruplex in 42.5% v/v CH₃CN/buffer solutions with the potassium phosphate buffer used previously. Samples were prepared at 6 μM strand concentration and annealed by placing in boiling water for 10 minutes, then slowly cooled to room temperature. Samples were allowed to cool slowly in the water bath overnight before use. The CD spectra are depicted below in Figure 5 - 9.
Curiously, the unmodified quadruplex has been forced into the parallel topology. The CD spectrum generated is in agreement with reports within the literature, displaying a large peak at 265 nm with a negative amplitude at 245 nm; the shoulder at 290 nm indicates some anti-parallel/hybrid topologies remain. The equilibrium of conformations exhibited by the mHTelo oligonucleotides in CH$_3$CN was not forced to the parallel conformation, although there were specific changes within the CD spectra generated upon the addition of acetonitrile.
Recall that $^{Fur}dG@3$ and $^{Fur}dG@10$ required an *anti* glycosidic orientation in the hybrid topologies; they are expected to prefer the anti-parallel topologies within K$^+$; they also exhibited lower degrees of stabilization during $T_m$ analysis (Table 5 - 2). The CD spectrum of $^{Fur}dG@3$ modified oligonucleotides was relatively unchanged upon addition of CH$_3$CN, suggesting that it remained in an anti-parallel conformation. The CD spectrum (Figure 5 - 9 (b)) features (1) a strong peak at 290 nm, indicating stacking of guanosines with different GBA; and (2) a positive peak at 240 nm coupled with a negative peak at 260 nm, which indicates the absence of stacking of guanosines with the same GBA. These findings are consistent with an anti-parallel G-quadruplex. The CD spectrum for $^{Fur}dG@10$ (Figure 5 - 9 (d)) displayed noticeable changes upon addition of CH$_3$CN. Specifically, the CD spectrum for CH$_3$CN shows the appearance of a negative amplitude band at 260 nm from what previously was a trough. The changed CD signal indicates the absence of guanosine stacking with the same GBA (positive signal at 240 nm, negative signal at 260 nm) which suggests that CH$_3$CN has induced formation of an anti-parallel topology.

Unlike $^{Fur}dG@3$ or $^{Fur}d@10$, the $^{Fur}dG@8$ modification showed significant $T_m$ stabilization in K$^+$, which suggests that G8 adopts a *syn* orientation in both hybrid and anti-parallel topologies. In this case, the hybrid forms were the most prevalent topology in K$^+$ solution. Analysis of the CD spectral changes upon introduction of CH$_3$CN suggests that the hybrid forms were further stabilized compared to the anti-parallel topology. The negative amplitude at 240 nm significantly grew in intensity and the trough that was present at 260 nm disappeared; this signal indicated the presence of guanosines stacking with the same GBA. The spectrum also contained a marginal increase at the amplitude at 290 nm, confirming that there is also guanosine stacking with different GBA. Altogether, this leads us to conclude that larger amounts of hybrid topologies being present in solution, as opposed to the parallel quadruplex which would not display an increased intensity at 290 nm. The CD spectrum for the $^{Fur}dG@8$
mHTelo in CH$_3$CN is very similar to that of the unmodified HTelo22 spectrum in K$^+$ which is known to favour the hybrid form quadruplexes.

In summary, addition of CH$_3$CN to induce dehydrating conditions did not promote the parallel topology in HTelo sequences. This was unexpected, since Fur$^\text{dG}$ can adopt an anti conformation within the mTBA oligonucleotides. The inability of mHTelo oligonucleotides to attain a parallel quadruplex may be explained by the orientation of the loop structures (Figure 5 - 2). Analysis of the parallel, propeller topologies crystal structure suggests that the loops may sterically clash with a modification at C8. In the parallel quadruplex, the loops project horizontally and wrap around the quadruplex structure, as opposed to the edgewise loops in the anti-parallel structures. At any position within the sequence, C8 projects directly toward either the phosphate backbone or the loop structure of the quadruplex, which likely destabilizes this topology and detracts from its formation.

Fluorescence spectra of the folded quadruplexes in CH$_3$CN/buffer can be seen below in Figure 5 - 10. Interestingly, they show increased emission intensities. This was not due to non-specific effects of CH$_3$CN, since no increase in intensity was observed when CH$_3$CN was added to prefolded quadruplexes in K$^+$ buffer. Furthermore, the quantum yield of the nucleoside adduct decreases from CH$_3$CN (0.244) compared to H$_2$O (0.46). Therefore, the observed change in emission must be due to a transition between topologies. The increased emission intensities in CH$_3$CN suggest that Fur$^\text{dG}@3$ and Fur$^\text{dG}@10$ mHTelo sequences attain larger amounts of anti-parallel topologies. The fact that Fur$^\text{dG}@8$ displayed larger emission intensities despite the CD spectrum indicating it attained more hybrid forms does not support this. It is hard to make direct relationships between emission intensity and an equilibrium of topologies as the Fur$^\text{dG}$ probe will be experiencing multiple environments.
Figure 5 - 10: Excitation and emission spectral overlays comparing various Fur\(^{\text{dG}}\) mHTelo22 oligonucleotides in K\(^{+}\) buffer (solid line) or 42.5% v/v CH\(_3\)CN/K\(^{+}\) buffer (dotted line). (a) Fur\(^{\text{dG@3}}\); (b) Fur\(^{\text{dG@8}}\); (c) Fur\(^{\text{dG@10}}\).
5.4.3.2. **Response to Conformational Selection**

5.4.3.2.1. **Addition of PEG**

To determine if the equilibrium can be forced to the parallel stranded quadruplex, PEG-600 (Polyethylene Glycol) was used. Studies by Buscaglia and coworkers found that interactions between PEG and the quadruplex seemed to be a strong driving force behind the conformational change from anti-parallel to parallel topologies. They reported differential binding of PEG to the propeller conformation, which shifts the equilibrium to the preferred form, describing a conformational selection mechanism as opposed to a molecular crowding or water activity mechanism. It is important to note that many reports within the literature describe PEG as a molecular crowding agent, and there is some debate to how it drives conformational changes within HTelo sequences. The effect of PEG was tested by making 6 μM quadruplex in 42.5% v/v PEG-600/buffer solutions exactly as described in Section 5.4.3.1. The CD spectra are depicted below in Figure 5 - 11.
The addition of PEG-600 forced the unmodified HTelo oligonucleotide equilibrium towards the parallel stranded topology. PEG-600 stabilizes the parallel quadruplex more effectively than CH$_3$CN, which suggests that its mechanism of action is different (supporting the conformational selection mechanism). The signal of the unmodified HTelo quadruplex in PEG solution is now typical of a parallel strand configuration.
The ability of various mHTelo sequences to attain the parallel structured quadruplex is dependent on the position of modification. Fur\textsubscript{dG}@8 most readily adopted a parallel conformation (Figure 5 - 11 (c)); this can be seen by the CD spectrum showing the characteristic positive signal at 260 nm and negative signal at 240 nm; a proportion of hybrid/anti-parallel topologies are still present as the peak at 290 nm remained. Fur\textsubscript{dG}@3 was also able to form a parallel quadruplex structure with the anti-parallel topology remaining (Figure 5 - 11 (b)). Fur\textsubscript{dG}@10 remained in a mixture of hybrid/anti-parallel type topologies. Addition of PEG to Fur\textsubscript{dG}@10 (Figure 5 - 11 (d)) resulted in a more intense signal at 295 nm and relatively little change anywhere else. This was not an expected result, as it indicates that PEG was not effective at inducing the formation of the parallel conformation for Fur\textsubscript{dG}@10. Increased amplitude in the CD spectrum at 295 nm is indicative of more guanosine stacking interactions with different GBAs, which could be a result of either hybrid/anti-parallel topologies and is ambiguous.

Probe performance was once again determined using fluorescence spectroscopy; the excitation and emission spectra generated for the folded quadruplexes in PEG-600 can be seen in Figure 5 - 12. These spectra once again show increased emission intensities. This was not due to any non-specific effect of PEG, since in a control titration with PEG-600 on prefolded quadruplexes in K\textsuperscript{+} buffer no increase in intensity occurred. Furthermore, the Fur\textsubscript{dG} nucleoside has been shown to be quenched in the presence of increased solvent viscosity, via the introduction of glycerol.\textsuperscript{89}
Figure 5 - 12: Excitation and emission spectral overlays comparing $^{\text{Fur}}$dG mHTelo22 oligonucleotides in K$^+$ buffer (solid line) or 42.5% v/v PEG-600/K$^+$ buffer (dotted line). (a) $^{\text{Fur}}$dG@3; (b) $^{\text{Fur}}$dG@8; (c) $^{\text{Fur}}$dG@10.

Upon transitioning from the hybrid-type to the parallel quadruplex, $^{\text{Fur}}$dG@3 exhibits moderately increased fluorescence intensity ($I_{\text{rel}} = 1.4$) and a strong peak at approximately 255 nm in the excitation spectrum; this peak even exceeds the one at 290 nm that was shown earlier to be diagnostic for anti-parallel quadruplex structures. Reports by Dumas on various vinyl-bridged pyridyl and phenyl-modified dG probes within HTelo23 sequences showed significant intensity at
approximately 260 nm in the excitation spectra despite CD, confirming that the hybrid topologies were maintained, no excitation peaks at 290 were observed.\textsuperscript{79}

\textsuperscript{Fur}\textsubscript{dG}@8 also depicts a more pronounced peak in the excitation spectrum at approximately 250 nm; however, the excitation spectrum is still predominantly displaying the 290 nm peak. Larger changes in emission intensity were noted for \textsuperscript{Fur}dG@8, as introduction of PEG resulted in an $I_{\text{rel}}$ of 2.65. \textsuperscript{Fur}dG@10 displayed the largest change in fluorescence intensity with the 290 nm peak much larger than the 255 nm. Interestingly, this adduct had the largest fluorescence response to the addition of PEG-600, exhibiting an $I_{\text{rel}}$ of 10.1.

Definitive characterization of the fluorescence response of \textsuperscript{Fur}dG in a parallel quadruplex cannot be made using the addition of PEG-600, as no mHTelo oligonucleotide obtained a single topology in solution.

\textbf{5.4.3.2.2. Addition of N-Methyl Mesoporphyrin IX}

Nicolaudis et al. have shown that N-methyl mesoporphyrin IX (NMM) (Figure 5 - 5) selectively stabilizes the parallel conformation of HTelo22.\textsuperscript{52,93} They report the crystal structure of an HTelo22 dimer capped on both 3' ends by NMM leading to a 1:1 binding stoichiometry and report that the N-methyl group is integral for selectivity.\textsuperscript{52} The parallel conformational selectivity was attributed to the terminal 3' G-tetrad being unobstructed by loop sequences, compared to the anti-parallel basket form, which has both terminal G-tetrads hindered by lateral edgewise or diagonal loops.\textsuperscript{93}

We reproduced the NMM binding experiment with HTelo containing \textsuperscript{Fur}dG at various positions to determine whether the probe can report on the formation of the parallel complex. 5 equivalents of NMM were added to these solutions and the quadruplexes were annealed as described in Section 5.4.3.1. The CD spectra are depicted below in Figure 5 - 13.
Figure 5 - 13: CD spectral changes upon addition of NMM to mHTelo oligonucleotides. Solid lines represent 100% K\(^+\)-buffer solutions while dotted lines correspond to solutions containing 5 equivalents NMM. (a) Unmodified sequence (b) \textsuperscript{Fur}dG@3 (c) \textsuperscript{Fur}dG@8 (d) \textsuperscript{Fur}dG@10

The unmodified sequence was clearly transformed from the hybrid topology to a parallel topology (Figure 5 - 13 (a)). All modified oligonucleotides failed to preferentially attain a parallel topology. Overall, the response of the mHTelo oligonucleotides to NMM resembles that to acetonitrile (Figure 5 - 9). This was unexpected, since CH\(_3\)CN is assumed to influence
quadruplex formation through dehydration, whereas NMM is supposed to bind selectively. Fur\textsubscript{dG}@3 seems to have transitioned from an anti-parallel topology to the hybrid forms. Fur\textsubscript{dG}@8 has developed a parallel component to the CD spectrum (peak appears at 265 nm and negative at 240 nm), but clearly NMM was not able to force the equilibrium exclusively to the parallel conformation (large peak remains at 295 nm). Fur\textsubscript{dG}@10 once again has not changed its overall topology, and exhibited a reduction in amplitude at the 290 nm peak as observed in CH\textsubscript{3}CN.

Probe performance was once again determined using fluorescence spectroscopy; the excitation and emission spectra generated for the folded quadruplexes with NMM are depicted in Figure 5 - 14. These spectra show quenched emission intensities as well as specific changes to the emission bands.
Figure 5 - 14: Excitation and emission spectral overlays comparing various Fur dG mHTelo22 oligonucleotides in K+ buffer (solid line) or K+ buffer with 5 equivalents NMM (dotted line). (a) Fur dG@3; (b) Fur dG@8; (c) Fur dG@10.

The reduction of emission intensities observed for all Fur dG mHTelo oligonucleotides subjected to NMM can be attributed to a resonance energy transfer (RET) mechanism (alternatively called fluorescence resonance energy transfer (FRET)). This process occurs whenever the emission spectrum of a fluorophore (donor, Fur dG), overlaps with the absorption
spectrum of another molecule (acceptor, NMM). An overlay depicting the emission spectrum of \textsuperscript{Fur}dG with the absorption spectrum of NMM is presented in Figure 5 - 15.

![Spectral overlay depicting \textsuperscript{Fur}dG emission (blue) and NMM excitation (red).](image)

Figure 5 - 15: Spectral overlay depicting \textsuperscript{Fur}dG emission (blue) and NMM excitation (red).

The acceptor does not need to be fluorescent, the extent of energy transfer is determined by the distance between the donor and acceptor, and the extent of spectral overlap. RET does not involve emission of light by the donor. RET is a dipole-dipole interaction, it occurs through space and over longer distances than other mechanisms of quenching such as collisional encounters which are subject to charge shielding.

Figure 5 - 16 shows a molecular orbital schematic for RET. The fluorophore initially has two electrons in the HOMO. Absorption of a photon results in excitation of one electron to the LUMO. When RET occurs, the electron in the excited donor (D\textsubscript{R}\textsuperscript{*}) returns to the ground state. Simultaneously, an electron in the acceptor (A\textsubscript{R}) goes into a higher excited-state orbital. If the acceptor is fluorescent it may then emit (if the acceptor is non-fluorescent the energy is dissipated as heat).
RET can be used to measure the distances between sites on macromolecules, as the extent of energy transfer is determined by the distance between the donor and acceptor, the relative orientation of the donor and acceptor transition dipole moments, and spectral overlap as previously mentioned. Energy transfer has been used as a “spectroscopic ruler” for measurements of distance between macromolecules\textsuperscript{96} DNA hybridization,\textsuperscript{97} and DNA folding into various quadruplexes.\textsuperscript{98,99} Typically, FRET is performed using dyes attached externally to nucleic acids through a linker. This scenario complicates quantitative interpretation of experiments because of dye diffusion and reorientation.\textsuperscript{100} Probes positioned inside the DNA structures, as a replacement for one of the natural bases, provides tighter orientation which gives the advantage of reporting from inside sites of interest.

In this case the orientation of $^\text{Fur}dG$ is relatively fixed within the quadruplex. However, since NMM occurs both bound and in solution, it is not possible to measure the distance between $^\text{Fur}dG$ and the bound NMM molecule. Moreover NMM exhibits a small absorption band at 315 nm Figure 5 - 15, which results in some direct excitation of the dye accounting for increased emission intensities of the NMM at 620 nm.
Addition of NMM to prefolded quadruplexes in 42.5% v/v CH₃CN/K⁺ buffer did result in spectral changes. One can see the gradual quenching and splitting of the emission intensity which starts at 380 nm and then separates into two resolved peaks at 366 and 423 nm (Figure 5 - 17).

Figure 5 - 17: Addition of NMM to 6 μM solutions of FurG₃ mHTelo quadruplexes. (a) FurG₃@3 (b) FurG₃@8 (c) FurG₃@10. Prefolded quadruplex in 42.5% v/v CH₃CN/K⁺ buffer (solid line) was subjected to 4.5 μM additions of NMM (dotted traces).
Interpretation for the observed changes in emission spectra is complex. It is likely that an excited state reaction is occurring, specific for the interaction of the Fur$dG$ probe with NMM, which results in the spectral changes of the emission band. Due to the lack of change in the absorption spectrum of Fur$dG$ (other than quenching) the interaction is assumed to only be occurring in the excited state; if such an assumption is true, than the appearance of the specific effect will depend on the rates of diffusion of the Fur$dG$ and NMM. These types of processes are usually studied by measurement of the time-resolved emission spectra, experiments for which were unable to be completed with the equipment available.

Another possible cause for the interesting emission response of Fur$dG$ is the formation of an internal charge-transfer (ICT) state, which results in longer-wavelength emission following excitation. The Fur$dG$ nucleoside did not behave as a D-A “push-pull” analogue, which was observed for the CNPh$dG$, Q$dG$ and Py$dG$ as previously discussed. However, because ICT emission is dependent on the electron-donating and accepting properties of the groups within or attached to the fluorophore system, ICT emission could be a result of NMM interacting with the Fur$dG$ probe or the quadruplex itself, by adding donating character (electronic density) to the system.

5.5. Conclusions

Overall, the probing potential of Fur$dG$ within the HTelo oligonucleotide has provided insight into the complex conformational equilibrium that exists for this sequence. As was shown for the various mTBA oligonucleotides, Fur$dG$ probes display large $I_{rel}$ values comparing the dsDNA and quadruplex structures formed in solution; for HTelo oligonucleotides excitation at 320 nm resulted in values between 5.6 and 30, while excitation at 290 nm resulted in values between 9.4 and 39 (Table 5 - 3). This immediately suggests that Fur$dG$ could gain use as a probe to monitor any duplex-quadruplex exchange reactions within HTelo oligonucleotides, and for the interaction of HTelo with small molecules that influence these reactions. Typically, substrates that induce quadruplex formation from duplex structures are considered detrimental due to the
unknown side effects that would occur within the genome. Therefore, \textsuperscript{Fur}dG modified HTelo oligonucleotides could provide insight when searching for small molecules that are specific G-quadruplex-inducing ligands.

The positional dependence of the \textsuperscript{Fur}dG modification on quadruplex formation was interesting. The \textit{syn} preference of \textsuperscript{Fur}dG clearly resulted in stabilization of quadruplexes that maintained this glycosidic orientation (the hybrid and anti-parallel topologies), although it was still evident that the HTelo oligonucleotide was able to obtain parallel conformations in solution (induction by PEG and NMM). The fact that the parallel topology was best stabilized via the conformational selection mechanism as opposed to osmotic stress (via CH\textsubscript{3}CN) is beneficial, as this suggests that \textsuperscript{Fur}dG modified oligonucleotides could gain use for \textit{in vitro} studies to identify small molecules that specifically induce parallel conformations. Although attainment of a parallel conformation was clearly impeded by the \textsuperscript{Fur}dG probe, \textsuperscript{Fur}dG did not inhibit its formation; with the proper experimental design, this can filter out molecules that do not have a high propensity to induce and stabilize a parallel conformation. It also seemed evident that the hybrid forms of \textsuperscript{Fur}dG mHTelo oligonucleotides were less emissive than the anti-parallel form (Figure 5-7 (b) vs. (c)), which can provide insight into which conformations are favoured under different conditions.

Another interesting aspect of this research concerns the possible effects of DNA damage occurring within the telomeres, particularly those of C8-aryl-dg adduct formation. The sequence is rich in G and, therefore, likely prone to DNA damage by electrophilic species. Indeed, telomere shortening occurs during oxidative and inflammatory stress with guanine as the major site of damage.\textsuperscript{101} Fleming and coworkers have elucidated the propensity for oxidative damage of dG residues within the HTelo G-quadruplex. They show that the damage is localized to the terminal tetrads, as opposed to the central tetrad of the quadruplex, although this varied somewhat with the method of oxidation.\textsuperscript{101} The effect of adduct formation on quadruplex formation will
likely vary with the position in the sequence. Within the family of 8-aryl-dG adducts presented, Fur-dG has the lowest barrier to rotation about its glycosidic bond. Therefore, most 8-aryl-dG adducts would be unable to attain an anti conformation required for certain quadruplex topologies (as evidenced by the mTBA data). This will clearly impact telomeric function; it has been hypothesized that different telomeric structures are prevalent or required for certain biological situations, DNA damage via 8-aryl-dG adduct formation could have major implications on such telomere interactions, which is a scenario that has not been extensively studied.

In summary, the Fur-dG modification of the HTelo oligonucleotide provides a diagnostic handle that can be exploited for various in vitro studies aimed at understanding the implications of G-quadruplex folding topologies on biological function. Specifically, it could be used to study small molecule ligands that induce specific quadruplex folds, monitor quadruplex folding or unfolding, compare mechanisms of repair and replication in the telomeric region compared to standard dsDNA and determine interactions with various proteins.
5.6. References


(39) Neidle, S. *FEBS J.* 2010, 277, 1118.


(41) Wang, Y.; Patel, D. J. *Structure* 1993, 1, 263.


278


Appendix A: General Experimental Methods
Materials. Unless otherwise noted, commercial compounds were used as received, and in general, were purchased from Sigma-Aldrich (St. Louis, MO). All nucleosides, including dmf-dA, ac-dC, dmf-dG, and dT were purchased from ChemGenes (Wilmington, MA), boronic acids from Frontier Scientific (Logan, UT), and triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt hydrate (TPPTS) from Alfa Aesar (Ward Hill, MA). All normal (unmodified) oligonucleotides were synthesized and purified by Sigma-Aldrich Ltd (Oakville, ON), including Nar1 (5'-CTCGGCGCCATC), and its complementary strands Nar1'(C) (5'-GATGGCGCCGAG), Nar1'(G) (5'-GATGGGGCCGAG), Nar1'(10mer) (5'-GATGGCGCCGAG) and Nar1'(THF) (5'-GATGG(THF)GCCGAG; TBA (5'-GGTTGTTGTGGTTGG), and its complementary strands TBA' (5'-CCAACCACACCAACC), TBA'(9mer) (5'-ACCACACCA), TBA'(10mer) (5'-CCACACACCAAC), TBA'(11mer) (5'-AACCACACCAA); and HTelo22 (5'-AGGGTTAGGGTTAGGG) and its complementary strand HTelo22' (5'-CCCTAACCCTAACCCTAACC). Triethylammonium acetate (TEAA) buffer was prepared from glacial acetic acid and triethylamine (Et$_3$N), and buffer pH adjusted using glacial acetic acid or Et$_3$N. Water used for buffers and spectroscopic solutions was obtained from a Milli-Q filtration system (18.2 MΩ).

Equipment. $^1$H, $^{13}$C and $^{31}$P NMR spectra were recorded on 300, 400, or 600 MHz spectrometers in either DMSO-d$_6$, CDCl$_3$, CD$_2$Cl$_2$, CD$_3$CN, Acetone-d$_6$ or CD$_3$OD referenced to either TMS (0 ppm) or the respective solvent. Unless specified all NMR experiments were carried out at room temperature and processed using TopSpin 2.1 software. pH measurements were taken at room temperature with an Accumet 910 pH meter with an Accumet pH Combination Electrode with stirring. UV-Vis measurements were recorded on a Cary 300-Bio UV-Visible Spectrophotometer equipped with a 6 × 6 Multicell Block Peltier, stirrer and temperature controller with Probe Series II. Fluorescence measurements were recorded on a Cary Eclipse Fluorescence Spectrophotometer equipped with a 1 × 4 Multicell Block Peltier, stirrer and temperature controller with Probe Series
II. Circular Dichroism measurements were recorded on a Jasco J-815 CD spectropolarimeter equipped with a 1 × 6 Multicell block thermal controller and a water circulator unit.

**Photophysical Measurements.** Unless otherwise noted, stock solutions of nucleoside adducts for use in spectroscopic measurements were prepared in DMSO, due to sparing solubility in other solvents, to a concentration of 4 mM. Spectroscopic solutions were prepared directly in the cuvette. Quartz glass cells from Hellma Analytics were used for all spectroscopic measurements. For analysis of all nucleosides, UV-Vis spectra were recorded using 100-QS cells with a light path of 10 mm, and fluorescence spectra recorded using 101-QS cells with a light path of 10 × 10 mm. For analysis of all oligonucleotides, UV-Vis spectra were recorded using 108.002-QS cells with a light path of 10 mm, and fluorescence spectra recorded using 108.002F-QS cells with a light path of 10 × 2 mm. All fluorescence titration spectra were recorded using 119.004F-QS with a 10 × 4 mm light path. All UV-Vis measurements were recorded at room temperature and fluorescence measurements were recorded at 10°C, unless otherwise noted, with stirring and baseline correction. Unless otherwise noted, UV-Vis spectra were recorded from 400 to 220 nm. Fluorescence excitation spectra were recorded at the emission wavelength (maximum) of the adduct, from 220 nm to 10 nm below the emission wavelength, while fluorescence emission spectra were recorded at the excitation wavelength (maximum), from 10 nm above the excitation wavelength to 600 nm. Unless otherwise noted, for all fluorescence measurements with nucleosides, excitation and emission slit widths were kept constant at 2.5 nm, while for all fluorescence measurements with oligonucleotides, excitation and emission slit widths were kept constant at 5 nm.

**Kinetic Measurements.** FurG, CNBrdG, QdG and PhdG samples were submitted to Dr. Katie Rankin at the University of Guelph for hydrolysis studies. Rates were determined by monitoring the appearance of the deglycosylated species at its absorbance maximum, or excitation and
emission maxima, respectively. The wavelength maximum of the deglycosylated adduct was determined separately via a scan measurement.

**Oligonucleotide Synthesis.** Oligonucleotide synthesis of T₂XTT, 5'-DMT-TXTT (X = Fur-dG) and 8-aryl-dG modified NarI(12) (5’-CTCGGCXCCCATCCCTACGAGC), NarI(22) (5’-CTCGGCXCCATCCCTACGAGC), TBA (5’-GGTTXXTXTGTTGG), HTelo22 (5’-AGXGTTAXGTTAGGTAGGG) (with X = Fur-dG, Ph-dG, CNPh-dG, Bth-dG, Q-dG or Py-dG) was carried out on a 1 µmol scale using an automatic DNA synthesizer. For all modified phosphoramidites with 5’-O-DMT protection, synthesis was carried out using standard β-cyanoethylphosphoramidite chemistry (unmodified phosphoramidites (bz-dA-CE, ac-dC-CE, dmf-dG-CE, and dT-CE), activator (0.25 M 5-(ethylthio)-1H-tetrazole in CH₃CN), oxidizing agent (0.02 M I₂ in THF/pyridine/H₂O, 70/20/10, v/v/v), deblock (3% DCA in dry DCM), cap A (THF/2,6-lutidine/acetic anhydride), cap B (methylimidazole in THF), and solid supports (5’-DMT-dC(Ac), or 1000 Å controlled pore glass (CPG), or 5’-DMT-dT 1000 Å CPG or 5’-DMT-dG(dmf) 1000 Å CPG) according to referenced protocols. For phosphoramidites with 5’-O-DMPx protection, standard DNA synthesis conditions and 5’-ODMT phosphoramidites were utilized prior to coupling with the modified phosphoramidite. At this point the synthesis was halted, and the normal deblock solution (3% DCA in dry DCM) was replaced with 0.5% DCA in dry DCM. The synthesis was then allowed to proceed to completion using 5’-O-DMPx phosphoramidites (modified (6a–c), and unmodified (7b, 8b, 9b), Scheme 1). Following synthesis, oligonucleotides were cleaved from the solid support and deprotected using 2 mL of 30% ammonium hydroxide solution at 55 °C for 12 h and purified by RP-HPLC.

**Oligonucleotide Purification.** The 8-aryl-G-modified NarI(12) and NarI(22) oligonucleotide solutions were first filtered using syringe filters (PVDF 0.20 µm) and concentrated under diminished pressure. Purification was performed using an HPLC instrument equipped with an autosampler, a diode array detector (monitored at 258 nm and λabs of the incorporated modified
nucleoside), fluorescence detector (monitored at λex and λem of the incorporated modified nucleoside), and autocollector. Separation was carried out at 50 °C using a 5 μm reversed-phase (RP) semipreparative C18 column (100 × 10 mm) with a flow rate of 3.5 mL/min, and various gradients of buffer B in buffer A (buffer A = 95:5 aqueous 50 mM TEAA, pH 7.2/acetonitrile; buffer B = 30:70 aqueous 50 mM TEAA, pH 7.2/acetonitrile). Collected DNA samples were lyophilized to dryness and redissolved in 18.2 MΩ.

**Oligonucleotide Quantification.** Quantification was performed with UV−vis measurement using ε260. Extinction coefficients were obtained from the following Web site: http://www.idtdna.com/analyzer/applications/oligoanalyzer. The 8-aryl-G modified oligonucleotides were assumed to have the same extinction coefficient as the natural NarI(12) (102 100 M⁻¹ cm⁻¹) and NarI(22) (185 700 M⁻¹ cm⁻¹) oligonucleotides, TBA (143 300 M⁻¹ cm⁻¹), TBA' (138 600 M⁻¹ cm⁻¹), TBA'(10) (92600 M⁻¹ cm⁻¹), HTelo22 (228500 M⁻¹ cm⁻¹), HTelo'(22) (193700 M⁻¹ cm⁻¹).

**Quantification of titrants.** Quantification of titrants was performed with UV−vis measurement using εmax and extinction coefficients found on Sigma Aldrich product information files. Bovine Thrombin (72 150 M⁻¹ cm⁻¹ @ 280 nm); Bovine Serum Albumin (BSA) (44309 M⁻¹ cm⁻¹ @ 279 nm); N-methylmesoporphyrin (NMM) (145 000 M⁻¹ cm⁻¹ @ 379 nm).

**MS Analysis.** MS experiments for DNA identification were conducted by Aaron Witham at the University of Guelph, Guelph ON, on a quadrupole ion trap SL spectrometer. Masses were acquired in the negative ionization mode with an electrospray ionization source. Oligonucleotide samples were prepared in 90% Milli-Q filtered water/10% methanol containing 0.1 mM ammonium acetate. Full scan MS spectra were obtained by direct infusion at a rate of 5−10 μL/min. High resolution mass spectra (HRMS) of all C8-aryl-dG adducts was recorded at the McMaster Regional Centre for Mass Spectrometry at McMaster University in Hamilton, Ontario,
Canada, on a Micromass Global Ultima Quadrupole Time of Flight Spectrometer, operating in electrospray ionization (ESI) detecting positive ions.

**Annealing of DNA.** Oligonucleotide samples (1 mL volume total) were prepared in 100 mM M\(^+\)-Phosphate, pH 7.0 with 100 mM MCl (M\(^+\) = Na\(^+\) or K\(^+\)), using equivalent amounts (6 μM) of the unmodified or 8-aryl-dG modified oligonucleotide and its complementary strand (for duplexes). Quadruplexes required only the unmodified or 8-aryl-dG modified oligonucleotide to be present in 6 μM concentrations. Samples were then heated at 90°C in a water bath for 5 min, allowed to slow cool to room temperature and placed in the fridge overnight prior to use.

**UV Melting:** All thermal melting temperatures (T\(_m\)) of oligonucleotides were measured using a Cary 300-Bio UV–vis spectrophotometer equipped with a 6 x 6 multicell Peltier block-heating unit using Hellma 114-QS 10 mm light path cells. The UV absorption (at 260 nm for duplexes and 295 nm for G-quadruplexes) was monitored as a function of temperature and consisted of forward-reverse scans from 10 to 90 °C at a heating rate of 1 °C/min, which was repeated five times. The T\(_m\) values were determined using hyperchromicity calculations provided in the Varian Thermal software.

**Circular Dichroism (CD):** Spectra were recorded on a Jasco J-815 CD spectropolarimeter equipped with a 1 x 6 Multicell block thermal controller and a water circulator unit. Measurements were carried out in 100 mM M\(^+\)-phosphate, pH 7.0 with 100 mM MCl (M = Na\(^+\) or K\(^+\)), using 6 μM of the unmodified or 8-aryl-G modified oligonucleotide and its complementary strand in equivalent amounts. Quartz glass cells (110-QS) with a light path of 1 mm were used for measurements. Spectra were collected at 10 °C between 200 and 400 nm, with a bandwidth of 1 nm and scanning speed at 100 nm/min. The spectra were the averages of five accumulations that were smoothed using the Jasco software.
Computational Details for experimental data generated by Purshotam Sharma, University of Lethbridge, Lethbridge, AB, Canada.

MD simulations were performed on the 12-mer oligonucleotide (5′–CTCG1G2CG3CCATC) containing the adducted nucleoside (Fur dG, PhdG or QdG) at the G3 position. Initial structures were built using both syn and anti conformations of the adducted nucleosides. For Fur dG and QdG, the syn and anti conformations were built using two different orientations of the asymmetric C8-ring system with respect to the nucleobase, which are related by an approximate 180° flip along the θ torsion angle. The starting DNA structure was built using the NAB module of AMBER 11, and the C8-modification was introduced at the G3 position using Gaussview. The RED.v.III.4 program was used to calculate the partial charges of the adducted nucleosides, while parameters for bonded terms were taken from the generalized AMBER force field (GAFF) according to atom types assigned using ANTECHAMBER 1.4. The parmbsc0 modification to the parmm99 force field was used for the natural nucleosides. Each DNA system was neutralized with 22 sodium ions and solvated with an 8 Å T1P3P truncated octahedral water box. Initial minimization of the water molecules and sodium ions was performed with the DNA fixed using a steepest descent algorithm for the first 500 steps, and a conjugate gradient algorithm for the next 500 steps. Subsequently, the entire system was minimized for 2500 steps using 1000 steps of steepest descent followed by 1500 steps of conjugate gradient minimization. The system was then heated from 0 to 300 K using Langevin Temperature equilibration during which the DNA was weakly restrained with a force constant of 10 kcal mol^{-1} Å^{-2} to avoid large fluctuations within the DNA, and simulated for 20 ps at constant volume. Subsequently, a 40 ns constant pressure MD simulation was carried out on each system using the PMEMD module of AMBER 11 or 12. A representative structure was chosen for analysis from each simulation by clustering the entire corresponding simulation with respect to the θ and χ dihedral angles of the adducts using the ptraj module of AMBER. Free energies were calculated from each simulation trajectories within the
MM-PBSA approximation.\textsuperscript{10} In these calculations, the molecular mechanics and solvation free energy terms were evaluated using the Poisson-Boltzmann approach and 1000 snapshots (1 snapshot every 20 ps from the simulation trajectory), while the entropy term was evaluated from normal mode calculations using 100 snapshots (1 snapshot every 200 ps).
Appendix B: Characterization of Synthesized Products by $^{1}$H, $^{13}$C, $^{31}$P NMR Spectroscopy.
Figure B - 1: $^1$H NMR spectrum of 8-quinoline-boronic acid in CD$_3$OD.

Figure B - 2: $^{13}$C NMR spectrum of 8-quinoline-boronic acid in CD$_3$OD.
Figure B - 3: $^1$H NMR spectrum of 8-(2''-furanyl)-2'-dG (1a) in DMSO-d$_6$.

Figure B - 4: $^{13}$C NMR spectrum of 8-(2''-furanyl)-2'-dG (1a) in DMSO-d$_6$. 

291
Figure B - 5: $^1$H NMR spectrum of 8-phenyl-2'-dG (1b) in DMSO-d$_6$.

Figure B - 6: $^{13}$C NMR spectrum of 8-phenyl-2'-dG (1b) in DMSO-d$_6$. 
Figure B - 7: $^1$H NMR spectrum of 8-(4''-cyanophenyl)-2'-dG (1c) in DMSO-d$_6$.

Figure B - 8: $^{13}$C NMR spectrum of 8-(4''-cyanophenyl)-2'-dG (1c) in DMSO-d$_6$. 
Figure B - 9: $^1$H NMR spectrum 8-(8''-quinolyl)-2'-dG (1d) in DMSO-d$_6$.

Figure B - 10: $^{13}$C NMR spectrum 8-(8''-quinolyl)-2'-dG (1d) in DMSO-d$_6$. 
Figure B - 11: $^1$H NMR spectrum of 8-(2''-benzothienyl)-2'-dG (1e) in DMSO-d$_6$.

Figure B - 12: $^{13}$C NMR spectrum of 8-(2''-benzothienyl)-2'-dG (1e) in DMSO-d$_6$. 
Figure B - 13: $^1$H NMR spectrum 8-(pyrenyl)-2'-dG (1f) in DMSO-$d_6$.

Figure B - 14: $^{13}$C NMR spectrum 8-(pyrenyl)-2'-dG (1f) in DMSO-$d_6$. 
Figure B - 15: $^1$H NMR spectrum of $N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (2a) in DMSO-d$_6$.

Figure B - 16: $^{13}$C NMR spectrum of $N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (2a) in DMSO-d$_6$. 

297
Figure B - 17: $^1$H NMR spectrum of $N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (2b) in DMSO-$d_6$.

Figure B - 18: $^{13}$C NMR spectrum of $N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (2b) in DMSO-$d_6$. 
Figure B - 19: $^1$H NMR spectrum of $N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (2c) in DMSO-$d_6$.

Figure B - 20: $^{13}$C NMR spectrum of $N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (2c) in DMSO-$d_6$. 
Figure B - 21: $^1$H NMR spectrum $N^2$-(dimethylformamidyl)-8-(8$^\prime$-quinolyl)-2$'$-dG (2d) in DMSO-$d_6$.

Figure B - 22: $^{13}$C NMR spectrum $N^2$-(Dimethylformamidyl)-8-(8$^\prime$-Quinolyl)-2$'$-dG (2d) in DMSO-$d_6$. 
Figure B - 23: $^1$H NMR spectrum of $N^2$-(dimethylformamidyl)-8-(2''-benzothenyl)-2'-dG (2e) in DMSO-$d_6$.

Figure B - 24: $^{13}$C NMR spectrum of $N^2$-(Dimethylformamidyl)-8-(2''-Benzothienyl)-2'-dG (2e) in DMSO-$d_6$. 
Figure B - 25: $^1$H NMR spectrum of $N^2$-(dimethylformamidyl)-8-(pyrenyl)-2$'$-dG (2f) in DMSO-$d_6$.

Figure B - 26: $^{13}$C NMR spectrum $N^2$-(dimethylformamidyl)-8-(pyrenyl)-2$'$-dG (2f) in DMSO-$d_6$. 
Figure B - 27: $^1$H NMR spectrum of 5′-O-DMT-N$^2$-(dimethylformamidyl)-8-(2″-furanyl)-2′-dG (3a) in CD$_3$CN.

Figure B - 28: $^{13}$C NMR spectrum of 5′-O-DMT-N$^2$-(dimethylformamidyl)-8-(2″-furanyl)-2′-dG (3a) in CD$_3$CN.
Figure B - 29: $^1$H NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (3b) in DMSO-$d_6$.

Figure B - 30: $^{13}$C NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (3b) in DMSO-$d_6$. 
Figure B - 31: $^1$H NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (3c) in CD$_2$Cl$_2$.

Figure B - 32: $^{13}$C NMR spectrum of 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dGe (3c) in CD$_2$Cl$_2$. 
Figure B - 33: $^1$H NMR spectrum 5'-O-DMT-$^N^2$-(dimethylformamidyl)-8-(8"-quinolyl)-2'-dG (3d) in DMSO-$d_6$.

Figure B - 34: $^{13}$C NMR spectrum 5'-O-DMT-$^N^2$-(dimethylformamidyl)-8-(8"-quinolyl)-2'-dG (3d) in DMSO-$d_6$. 
306
Figure B - 35: $^1$H NMR spectrum of 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (3e) in CDCl$_3$.

Figure B - 36: $^{13}$C NMR spectrum of 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(2''-Benzothienyl)-2''-dG (3e) in CDCl$_3$. 
Figure B - 37: $^1$H NMR spectrum 5′-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2′-dG (3f) in CDCl$_3$.

Figure B - 38: $^{13}$C NMR spectrum 5′-O-DMT-$N^2$-(dimethylformamidyl)-8-(pyrenyl)-2′-dG (3f) in CDCl$_3$. 
Figure B - 39: $^1$H NMR spectrum DMPx-OMe in CDCl$_3$.

Figure B - 40: $^{13}$C NMR spectrum DMPx-OMe in CDCl$_3$. 
Figure B - 41: $^1$H NMR spectrum DMPx-Cl in CD$_2$Cl$_2$.

Figure B - 42: $^{13}$C NMR spectrum DMPx-Cl in CD$_2$Cl$_2$. 
Figure B - 43: $^1$H NMR spectrum of $5'$-O-DMPx-$N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (5a) in CD$_2$Cl$_2$.

Figure B - 44: $^{13}$C NMR spectrum of $5'$-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (5a) in CD$_2$Cl$_2$. 

311
Figure B - 45: $^1$H NMR spectrum of 5$'$-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2$'$-dG (5b) in CD$_2$Cl$_2$.

Figure B - 46: $^1$H NMR spectrum of 5$'$-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2$'$-dG (5b) in CD$_2$Cl$_2$. 
Figure B - 47: $^1$H NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (5c) in DMSO-$d_6$.

Figure B - 48: $^{13}$C NMR spectrum of 5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (5c) in DMSO-$d_6$. 
Figure B - 49: $^1$H NMR spectrum 5'-O-DMPx-$N^2$-(dimethylformamidyl)-dG (5g) in CDCl$_3$.

Figure B - 50: $^{13}$C NMR spectrum 5'-O-DMPx-$N^2$-(dimethylformamidyl)-dG (5g) in CDCl$_3$. 
Figure B - 51: $^1$H NMR spectrum $5'$-O-DMPx-$N^4$-(acetyl)-2'-dC (5h) in CD$_2$Cl$_2$.

Figure B - 52: $^{13}$C NMR spectrum of $5'$-O-DMPx-$N^4$-(acetyl)-2'-dC (5h) in CD$_2$Cl$_2$. 
Figure B - 53: $^1$H NMR spectrum 5′-O-DMPx-2′-dT (5i) in DMSO-d$_6$.

Figure B - 54: $^{13}$C NMR spectrum 5′-O-DMPx-2′-dT (5i) in DMSO-d$_6$. 

316
Figure B - 55: $^1$H NMR spectrum $5'$-O-DMPx-$N^2$-(dimethylformamidyl)-dG (5j) in CDCl$_3$.

Figure B - 56: $^{13}$C NMR spectrum $5'$-O-DMPx-$N^2$-(dimethylformamidyl)-dG (5j) in CDCl$_3$. 
Figure B - 57: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (4a) in CDCl$_3$.

Figure B - 58: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (4a) in CDCl$_3$. 
Figure B - 59: \( ^{31}P \) NMR spectrum of 3′-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5′-O-DMT-N\(^2\)-(dimethylform-amidyl)-8-(2′″-furanyl)-2′-dG (4a) in CDCl\(_3\).
Figure B - 60: $^1$H NMR spectrum of 3’-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5’-O-DMPx-N(2′)-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (6a) in CD$_2$Cl$_2$.

Figure B - 61: $^{13}$C NMR spectrum of 3’-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5’-O-DMPx-N(2′)-(dimethylformamidyl)-8-(2′′-furanyl)-2′-dG (6a) in CD$_2$Cl$_2$. 
Figure B - 62: $^{31}$P NMR spectrum of $3'^{-}O$-[(2-cyanoethoxy)(diisopropylamino)phosphino]-$5'^{-}O$-DMPx-$N^2$-(dimethylformamidyl)-8-(2''-furanyl)-2'-dG (6a) in CD$_2$Cl$_2$.
Figure B - 63: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (4b) in CDCl$_3$.

Figure B - 64: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (4b) in CDCl$_3$. 
Figure B - 65: $^{31}$P NMR spectrum of $3'$-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5$'$-O-DMT-$N^2$-(dimethylformamidyl)-8-phenyl-2$'$-dG (5b) in CDCl$_3$. 
Figure B - 66: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^2$-(dimethylformamidyl)-8-phenyl-2'-dG (6b) in CD$_2$Cl$_2$.

Figure B - 67: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-N$^2$-(dimethylformamidyl)-8-phenyl-2'-dG (6b) in CD$_2$Cl$_2$. 
Figure B - 68: $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMPx-$N^2$-(dimethylformamidyl)-8-phenyl-2'-dG (6b) in CD$_2$Cl$_2$. 
Figure B - 69: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (4c) in CDCl$_3$.

Figure B - 70: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (4c) in CDCl$_3$. 326
Figure B - 71: $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (4c) in CDCl$_3$.
Figure B - 72: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (6c) in CD$_3$OD.

Figure B - 73: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (6c) in CD$_3$OD.
Figure B - 74: $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-8-(4''-cyanophenyl)-2'-dG (6c) in CD$_3$OD.
Figure B - 75: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT N²-(dimethylformamidyl)-8-(8''-quinolyl)-2'-dG (4d) in DMSO-d$_6$.

Figure B - 76: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT N²-(dimethylformamidyl)-8-(8''-quinolyl)-2'-dG (4d) in CDCl$_3$. 
Figure B - 31P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-\(N^2\)-(dimethylformamidyl)-8-(8'-quinoyl)-2'-dG (4d) in DMSO-d6.
Figure B - 78: $^1$H NMR spectrum of $3'$-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- $5'$-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (4e) in Acetone-$d_6$.

Figure B - 79: $^{13}$C NMR spectrum of $3'$-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- $5'$-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (4e) in Acetone-$d_6$. 
Figure B - 80: $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT-$N^2$-(dimethylformamidyl)-8-(2''-benzothienyl)-2'-dG (4e) in Acetone-$d_6$. 
Figure B - 81: $^1$H NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (4f) in CD$_2$Cl$_2$.

Figure B - 82: $^{13}$C NMR of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG(4f) in CD$_2$Cl$_2$. 
Figure B - $^{31}$P NMR of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMT-N$^2$-(dimethylformamidyl)-8-(pyrenyl)-2'-dG (4f) in CD$_2$Cl$_2$. 
Figure B - 84: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dG in (6g) CD$_2$Cl$_2$.

Figure B - 85: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dG (6g) in CD$_2$Cl$_2$. 
Figure B - 86: $^{31}$P NMR spectrum of 3’-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5’-O- DMPx-$N^2$-(dimethylformamidyl)-2’-dG (6g) in CD$_2$Cl$_2$. 
Figure B - 87: $^1$H NMR spectrum $3'-O-[(2$-cyanoethoxy)(diisopropylamino)phosphino]-$5'$-$O$-DMPx-$N^{4}$-(acetyl)-2'-dC (6h) in CDCl$_3$.

Figure B - 88: $^{13}$C NMR spectrum $3'$-$O-[(2$-cyanoethoxy)(diisopropylamino)phosphino]-$5'$-$O$-DMPx-$N^{4}$-(acetyl)-2'-dC (6h) in CDCl$_3$. 
Figure B - 89: $^{31}$P NMR spectrum 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'-O-
DMPx-$N^\prime$-(acetyl)-2'-dC (6h) in CDCl$_3$. 
Figure B - 90: $^1$H NMR spectrum $3'$-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'$'$-O-DMPx-2'-dT (6i) in CD$_2$Cl$_2$.

Figure B - 91: $^{13}$C NMR spectrum $3'$-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]-5'$'$-O-DMPx-2'-dT (6i) in CD$_2$Cl$_2$. 
Figure B - 92: $^{31}$P NMR spectrum $3'$-$O$-[(2-cyanoethoxy)(diisopropylamino)phosphino]-$5'$-$O$-DMPx-$2'$-dT (6i) in CD$_2$Cl$_2$
Figure B - 93: $^1$H NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$.

Figure B - 94: $^{13}$C NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$. 
Figure B - 95: $^{31}$P NMR spectrum of 3'-O-[(2-cyanoethoxy)(diisopropylamino)phosphino]- 5'-O-DMPx-N$^2$-(dimethylformamidyl)-2'-dA (6j) in CD$_2$Cl$_2$. 
Appendix C: Characterization of C₈-Aryl-G Modified Oligonucleotides by ESI Mass Spectrometry and UV Spectroscopy
Figure C - 1: Mass spectrum of 12mer $\textit{NarI}^{\text{Fur}}\text{dG}$, obtained with an ESI source operated in negative mode.

Figure C - 2: Mass spectrum of 12mer $\textit{NarI}^{\text{Bth}}\text{dG}$, obtained with an ESI source operated in negative mode
Figure C - 3: Mass spectrum of 12mer NarI $^{\text{CNPh}}$dG, obtained with an ESI source operated in negative mode.

Figure C - 4: Mass spectrum of 12mer NarI $^{\text{Q}}$dG, obtained with an ESI source operated in negative mode.
Figure C - 5: Mass spectrum of 12mer NarI PydG, obtained with an ESI source operated in negative mode.

Figure C - 6: Mass spectrum of 12mer NarI PhdG, obtained with an ESI source operated in negative mode.
Figure C - 7: Mass spectrum of 22-mer\textit{ Narf}^\text{Fur}dG, obtained with an ESI source operated in negative mode.

Figure C - 8: Mass spectrum of 22-mer\textit{ Narf}^\text{Bth}dG, obtained with an ESI source operated in negative mode.
Figure C - 9: Mass spectrum of 22-mer NarI QdG, obtained with an ESI source operated in negative mode.

Figure C - 10: Mass spectrum of 22-mer NarI PydG, obtained with an ESI source operated in negative mode.
Figure C - 11: Mass spectrum of 22-mer NarI PhdG, obtained with an ESI source operated in negative mode.

Figure C - 12: Mass spectrum of TBA Fur dG, obtained with an ESI source operated in negative mode.
Figure C - 13: Mass spectrum of TBA[^B][dG], obtained with an ESI source operated in negative mode.

Figure C - 14: Mass spectrum of TBA[^CNPh][dG], obtained with an ESI source operated in negative mode.
Figure C - 15: Mass spectrum of TBA$^0$dG, obtained with an ESI source operated in negative mode.

Figure C - 16: Mass spectrum of TBA$^\text{Py}$dG, obtained with an ESI source operated in negative mode.
Figure C - 17: Mass spectrum of TBA$_2$Fur$_n$dG + Py$_n$dG, obtained with an ESI source operated in negative mode.

Figure C - 18: Mass spectrum of TBA$_2$Fur$_n$dG + CN$_n$Ph$_n$dG, obtained with an ESI source operated in negative mode.
Figure C - 19: Mass spectrum of HTelo22 FurG obtained with an ESI source operated in negative mode.