

**TEMPLATE-ASSEMBLED SYNTHETIC G-QUARTETS (TASQS)**

**by**

**MEHRAN NIKAN**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF**

**DOCTOR OF PHILOSOPHY**

**in**

**THE FACULTY OF GRADUATE STUDIES**

**(Chemistry)**

**THE UNIVERSITY OF BRITISH COLUMBIA**

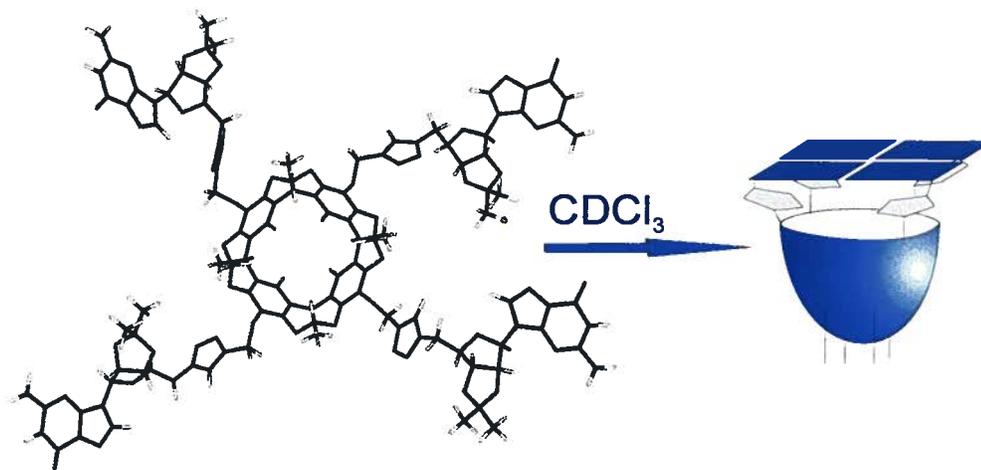
**November 2008**

**© Mehran Nikan, 2008**

## Abstract

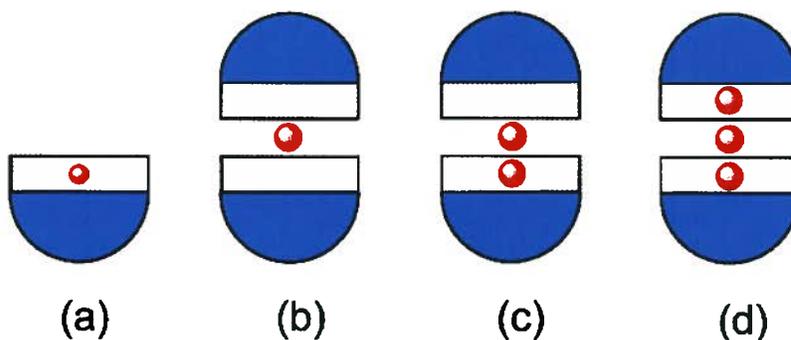
Fabrication of functional supramolecular structures requires a certain degree of control which may not be achieved by relying solely on noncovalent interactions. The current study aims to investigate the effect of a rigid cavitant template on morphology, function and stability of lipophilic G-quadruplexes. The first Chapter of this thesis introduces different aspects of G-quadruplex chemistry and explains how these structures are particularly suited for the creation of supramolecular architectures.

The second Chapter of this thesis presents the synthesis and self-assembly of a new class of supramolecular architectures composed of four guanosines attached to a rigid cavitant template. These structures, named template-assembled synthetic G-quartets (TASQs), were synthesized via the “click” reaction and manifest an ordered topology dictated by the template. The lipophilic TASQs were found to self-associate spontaneously to form a singular basket-like structure in chloroform. Moreover, it was found that TASQs form cation-free G-quartets which exhibit remarkable stability under this condition.



**TASQs self-associate in chloroform**

The third Chapter of this thesis describes the preparation, characterization and solution study of the cation-bound complexes TASQ·Na<sup>+</sup>, TASQ·K<sup>+</sup>, TASQ·Cs<sup>+</sup>, and TASQ·Sr<sup>2+</sup>. Cations play a major role in controlling the morphology and stability of G-quadruplexes. The analysis of the cation-specific structures of TASQs reveals the formation of a monomeric G-quartet for Na<sup>+</sup> and Sr<sup>2+</sup>, a dimeric system for Cs<sup>+</sup> and a mixture of monomers and dimers for K<sup>+</sup>. The factors governing the formation of these structures were evaluated, the selectivities of TASQs for cations were determined, and the cation-dependent structural transformations were studied.



**Polymorphism of TASQ-cation assemblies: Only structures consistent with (a) and (c) were observed**

The fourth Chapter describes the efforts towards synthesizing a hydrophilic TASQ via the “click” reaction. The following steps have been taken: 1) a water-soluble cavitand has been successfully synthesized and characterized, which can potentially serve as a hydrophilic template, and 2) two oligonucleotides have been appropriately functionalized and preliminary coupling reactions were attempted. The next phases of this research along with potential future directions are discussed in Chapter five.

## Table of Contents

<b>Abstract</b> .....	ii
<b>Table of Contents</b> .....	iv
<b>List of Tables</b> .....	ix
<b>List of Figures</b> .....	x
<b>List of Schemes</b> .....	xv
<b>List of Abbreviations</b> .....	xvi
<b>Acknowledgements</b> .....	xviii
<b>Co-Authorship</b> .....	xix
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>1</b>
1.1 Background.....	1
1.2 Thesis Goals.....	2
1.3 Thesis Overview.....	2
1.4 G-Quartet in Supramolecular Chemistry and Biology.....	3
1.4.1 Properties of Nucleotides.....	3
1.4.2 Guanine Self-Assembly: G-Quartet vs. G-Ribbon.....	5
1.4.3 Nucleic Acid G-Quadruplexes.....	6
1.4.4 Lipophilic G-Quadruplexes.....	8
1.4.4.1 Structure of Lipophilic G-Quadruplexes.....	9
1.4.5 The Role of Cations.....	11
1.4.6 Cation-Free G-Quartets.....	12
1.4.7 A G-Quartet with Extra H-Bonds.....	14
1.4.8 G-Quartet in Material Science and Nanotechnology.....	15
1.4.8.1 Liquid Crystalline Phases.....	16
1.4.8.2 Supramolecular Polymers.....	17
1.4.8.3 Nanoparticle Assembly.....	18
1.4.8.4 Molecular Electronics.....	19
1.4.8.5 Nanowires and G-Wires.....	20
1.4.8.6 Nanomachines.....	21
1.4.9 G-Quartet and Chiral Resolution.....	23
1.4.10 G-Quartet and Equilibriums.....	25
1.4.10.1 Dynamic Covalent Chemistry (DCC).....	25

1.4.10.2 Dynamic Cation Binding and Release .....	29
1.4.11 Nucleic Acid G-Quadruplexes as Probes.....	30
1.4.11.1 G-Quadruplexes as Cation-Sensors.....	30
1.4.11.2 G-Quadruplexes as Protein-Sensors.....	33
1.4.11.3 G-Quadruplexes as Oligonucleotide-Sensors .....	35
1.4.12 G-Quadruplexes as Catalysts .....	36
1.4.13 Miscellaneous G-Quadruplex-Based Structures .....	39
1.4.13.1 Peptide Nucleic Acid (PNA) G-Quadruplex.....	39
1.4.13.2 Bunch-Oligonucleotides.....	40
1.5 Summary and Conclusions .....	42
1.6 References.....	43
<b>CHAPTER 2:       Synthesis, Characterization, and Solution Studies of Lipophilic                       Template-Assembled Synthetic G-Quartets (TASQs).....</b>	<b>51</b>
2.1 Introduction.....	51
2.2 Rationale Behind Thesis .....	51
2.3 Results and Discussion .....	52
2.3.1 TASQ via “Click” Reaction.....	52
2.3.2 Synthesis of Cavitands <b>6a-c</b> .....	54
2.3.3 Synthesis of 5'-Azido-2',3'- <i>O</i> -Isopropylidene Guanosine <b>9</b> .....	55
2.3.4 Synthesis of TASQs <b>10a-c</b> .....	55
2.3.5 The Signal Assignment Strategy.....	56
2.3.5.1 Signal Assignment .....	58
2.3.6 Regiochemistry .....	64
2.3.7 Solution Structure Determination. ....	65
2.3.8 Variable-Temperature <sup>1</sup> H NMR Spectroscopic Studies .....	68
2.3.9 Infrared Spectroscopic Studies .....	72
2.3.10 CD Spectroscopic Studies.....	74
2.3.11 Diffusion NMR Studies .....	75
2.3.11.1 Translational Self-Diffusion.....	76
2.3.11.2 The Theory of Pulse Field-Gradient (PFG) NMR Spectroscopy .....	77
2.3.11.3 The Size Determination of TASQ <b>10c</b> by PFG NMR Spectroscopy .....	80
2.3.12 The Role of Water.....	83
2.4 Summary and Conclusions .....	85
2.5 Experimental .....	86

2.5.1	General	86
2.5.2	Synthesis of Cavitands <b>6a-c</b>	87
2.5.3	Synthesis of 5'-Azido-2',3'- <i>O</i> -Isopropylidene Guanosine <b>9</b>	88
2.5.4	Synthesis of TASQs <b>10a-c</b>	88
2.5.5	Synthesis of 5'-Azido-2',3'- <i>O</i> -Isopropylidene Adenosine <b>15</b>	90
2.5.6	Synthesis of Compound <b>16c</b>	92
2.5.7	PFG NMR Spectroscopic Experiments	93
2.5.7.1	General	93
2.5.7.2	Experimental (Bruker) Parameters for PFG NMR of <b>10c-16c</b> in CDCl <sub>3</sub>	93
2.5.7.3	Experimental (Bruker) Parameters for PFG NMR of <b>10c-16c</b> in DMSO-d <sub>6</sub>	94
2.6	References	95
<b>CHAPTER 3:</b>	<b>Cation-Complexation Behavior of TASQs: The Preparation and Structural Characterization of TASQ·Cation Assemblies</b>	<b>98</b>
3.1	Introduction	98
3.2	Rationale for the Study of TASQ·Cation Assemblies	98
3.3	Results and Discussion	100
3.3.1	Synthesis of TASQs	100
3.3.2	<sup>1</sup> H NMR Spectroscopic Studies	100
3.3.3	TASQ <b>10c</b> ·Na <sup>+</sup>	102
3.3.3.1	The Size Determination of TASQ <b>10c</b> ·Na <sup>+</sup> by PFG NMR Spectroscopy	105
3.3.4	TASQ <b>10c</b> ·Sr <sup>2+</sup>	107
3.3.4.1	The Size Determination of TASQ <b>10c</b> ·Sr <sup>2+</sup> by PFG NMR Spectroscopy	110
3.3.5	TASQ <b>10c</b> ·K <sup>+</sup>	111
3.3.6	TASQ <b>10c</b> ·Cs <sup>+</sup>	114
3.3.6.1	The Size Determination of TASQ <b>10c</b> ·Cs <sup>+</sup> by PFG NMR Spectroscopy	115
3.3.7	The Anion Effect	119
3.3.8	Gas-Phase Structures	120
3.3.9	Selectivities of TASQs	122
3.3.10	Interconversion of TASQ <b>10c</b> ·Na <sup>+</sup> to TASQ <b>10c</b>	125
3.3.11	CD Spectroscopy	127
3.4	Summary and Conclusions	128
3.5	Experimental	130
3.5.1	General	130
3.5.2	Picrate Extraction Experiments	131
3.5.3	PFG NMR Spectroscopic Experiments	132

3.5.3.1 Experimental (Bruker) Parameters for PFG NMR Experiments .....	132
3.5.4 Atomic Absorption Experiments .....	133
3.6 References.....	134
<b>CHAPTER 4:      Towards Hydrophilic TASQs.....</b>	<b>138</b>
4.1 Introduction.....	138
4.2 Rationale for the Study of Hydrophilic TASQs.....	138
4.3 Results and Discussion .....	139
4.3.1 The Oligonucleotide Approach.....	139
4.3.1.1 Solid-Phase Functionalization of the Oligonucleotides .....	141
4.3.1.2 The Coupling Reaction of 5'-Azido Oligonucleotides <b>20a-b</b> .....	142
4.3.2 The Water-Soluble Cavitand Approach.....	147
4.3.2.1 Synthesis of Phosphate-Footed Tetrakis ( <i>O</i> -Propargyl) Cavitand <b>30</b> .....	148
4.3.3 The Phosphoramidite Approach.....	151
4.4 Conclusions.....	155
4.5. Experimental .....	157
4.5.1 Phosphate-Footed Tetrakis ( <i>O</i> -Propargyl) Cavitand Synthesis .....	157
4.5.1.1 General .....	157
4.5.1.2 Synthesis of Dodecol <b>22</b> .....	158
4.5.1.3 Synthesis of Tetrabromo Dodecol <b>23</b> .....	158
4.5.1.4 Synthesis of Hydroxyl-Footed Tetrabromo Cavitand <b>24</b> .....	159
4.5.1.5 Synthesis of TBDMS-Protected Tetrabromo Cavitand <b>25</b> .....	160
4.5.1.6 Synthesis of TBDMS-Protected Tetrol <b>26</b> .....	161
4.5.1.7 Synthesis of TBDMS-Protected Tetrakis ( <i>O</i> -Propargyl) Cavitand <b>27</b> .....	161
4.5.1.8 Synthesis of Hydroxyl-Footed Tetrakis ( <i>O</i> -Propargyl) Cavitand <b>28</b> .....	162
4.5.1.9 Synthesis of <i>t</i> -Butyl Phosphorylated Tetrakis ( <i>O</i> -Propargyl) Cavitand <b>29</b> ....	163
4.5.1.10 Synthesis of Phosphate-Footed Tetrakis ( <i>O</i> -Propargyl) Cavitand <b>30</b> .....	164
4.5.2 Oligonucleotide Synthesis .....	165
4.5.2.1 General .....	165
4.5.2.2 Synthesis of 5'-Iodo Oligonucleotides <b>18a-b</b> .....	166
4.5.2.3 Synthesis of 5'-Azido Oligonucleotides <b>20a-b</b> .....	166
4.5.3 Phosphoramidite Synthesis .....	170
4.5.3.1 General .....	170
4.5.3.2 Synthesis of Tetrakisphosphoramidite Cavitand <b>34</b> .....	170
4.5.3.3 Synthesis of <i>N</i> <sup>2</sup> -Isobutyryl-2',3'- <i>O</i> -Isopropylidene Guanosine <b>33</b> .....	172

4.5.3.4 Synthesis of <i>N</i> <sup>2</sup> -Isobutyryl-2',3'- <i>O</i> -Isopropylidene Guanosine Phosphoramidite <b>36</b> .....	173
4.6 References.....	174
<b>CHAPTER 5: Summary, and Future Work.....</b>	<b>177</b>
5.1 Thesis Summary.....	177
5.2 Future Work.....	179
5.2.1 Lipophilic TASQ Project.....	180
5.2.1.1 Complementary Work.....	180
5.2.1.1 Future Applications.....	180
5.2.2 Hydrophilic TASQ Project.....	183
5.2.2.1 Complementary Work.....	183
5.2.2.2 Future Applications.....	186
5.3 References.....	187

## List of Tables

<b>Table 2.1.</b>	Spectral assignments of <b>10c</b> in DMSO-d <sub>6</sub> and CDCl <sub>3</sub> at ambient temperature. ...	59
<b>Table 2.2.</b>	Table of chemical shifts for a solution of <b>10c</b> in CDCl <sub>3</sub> at four different temperatures. ....	72
<b>Table 2.3.</b>	Diffusion coefficients (D) of <b>10c</b> and <b>16c</b> in CDCl <sub>3</sub> and DMSO-d <sub>6</sub> at 295 K.....	81
<b>Table 3.1.</b>	Diffusion coefficients (D) of <b>10c</b> ·Na <sup>+</sup> and <b>16c</b> in CDCl <sub>3</sub> at 295 K. ....	105
<b>Table 3.2.</b>	Diffusion coefficients (D) of <b>10c</b> ·Sr <sup>2+</sup> and <b>16c</b> in CDCl <sub>3</sub> at 295 K. ....	110
<b>Table 3.3.</b>	Diffusion coefficients (D) of <b>10c</b> ·Cs <sup>+</sup> and <b>10c</b> in CDCl <sub>3</sub> at 295 K. ....	117
<b>Table 3.4.</b>	The selectivities of TASQs. The magnitude of these selectivities is estimated as more than a factor of 20 between each pair of cations.....	124
<b>Table 3.5.</b>	Na content (μg/g) of TASQ samples.. ....	133
<b>Table 4.1.</b>	Reaction conditions examined for the coupling <sup>a</sup> of <b>20a</b> to <b>6a</b> .....	143

## List of Figures

<b>Figure 1.1.</b>	A tetrol cavitand.....	1
<b>Figure 1.2.</b>	Two sample nucleotides of guanine and cytosine. The numbering schemes for the base and sugar subunits are also presented. ....	3
<b>Figure 1.3.</b>	The purine (adenine, guanine) and pyrimidine (cytosine, thymine, uracil) bases of nucleic acids. ....	4
<b>Figure 1.4.</b>	The <i>syn</i> and <i>anti</i> conformations of guanosine. ....	4
<b>Figure 1.5.</b>	Watson-Crick and Hoogsteen base-pairing of nucleobases. Cytosine has to be protonated to be able to participate in Hoogsteen pairing with guanine or another molecule of cytosine. This mode of H-bonding is important in the formation of triplexes and i-motifs. ....	5
<b>Figure 1.6.</b>	A schematic representation of a) G-quartet, b) G-quadruplex (the G-quartet units are rotated about 30° in respect to each other). Images (c) and (d) depict G-ribbons (arrows indicate dipole moments).....	6
<b>Figure 1.7.</b>	a) Tetramolecular, b,c) bimolecular and d) unimolecular G-quadruplexes (arrows indicate 3' → 5' polarity). ....	8
<b>Figure 1.8.</b>	A schematic representation of an octamer formed from 3',5'-didecanoyl-2'-deoxyguanosine.....	9
<b>Figure 1.9.</b>	A schematic representation of the crystal structure of a lipophilic guanosine hexadecamer formed in the presence of a) K <sup>+</sup> and Cs <sup>+</sup> picrates, b) M <sup>2+</sup> picrates. The H-bonding of the picrate anion to the quartets is also presented. ....	10
<b>Figure 1.10.</b>	The <i>syn</i> conformation adopted by 8-(4- <i>N,N</i> -dimethylamino) guanosine directs the formation of a cation-free G-quartet. ....	12
<b>Figure 1.11.</b>	a) Calix[4]arene-guanosine conjugate, and b) calix[4]arene-guanosine dimers..	13
<b>Figure 1.12.</b>	The network of cation-free G-quartet made from guanine on gold surfaces. The H-bonds that connect the adjacent G-quartets are marked with circles. ....	14
<b>Figure 1.13.</b>	The G-quartet obtained from 8-aryl-2'-deoxyguanosine. The extra H-bonds are circled. ....	15
<b>Figure 1.14.</b>	A schematic representation of the liquid crystalline phases obtained from lipophilic or hydrophilic guanosine compounds. Depending on concentration, ions, and temperature, either the cholesteric or the hexagonal mesophases is obtained. ....	16
<b>Figure 1.15.</b>	The self-assembly of bisiminoboronate-guanosine to a G-quartet-based polymeric film in the presence of K <sup>+</sup> . The B-N dative bond of the monomer helps rigidify the polymer and also prevents the hydrolysis of the reversible boronate-guanosine ester bond. ....	17
<b>Figure 1.16.</b>	a) 3'-thiolpropyl deoxyguanosine phosphate and b) gold nanoparticle assembly using G-quartet formation (cations are omitted for clarity). ....	18
<b>Figure 1.17.</b>	A schematic representation of the simple transistor made from G-ribbons.....	19
<b>Figure 1.18.</b>	A schematic representation of the competing structures which can be adopted by telomeric DNA d(G <sub>4</sub> T <sub>2</sub> G <sub>4</sub> ). The experimental conditions can be adjusted to favor the formation of G-wires. As shown below the formation of a G-wire is the result of the interaction between the 3' end of one helix and the 5' end of another helix (cations are omitted for clarity). ....	20

<b>Figure 1.19.</b>	A schematic representation of a G-quadruplex-based nanomachine. The quadruplex-forming strand extends into a duplex upon addition of a complementary strand (C-fuel) and shrinks back to G-quadruplex via addition of G-fuel (cations are omitted for clarity). .....	21
<b>Figure 1.20.</b>	Copper-driven transition of a G-quadruplex to a random coil and vice versa.....	22
<b>Figure 1.21.</b>	Homochiral and heterochiral G-quadruplexes obtained from extraction of ( <i>D, L</i> )-5'-silyl-2',3'-isopropylidene guanosine with Ba <sup>2+</sup> and K <sup>+</sup> picrates. ....	23
<b>Figure 1.22.</b>	a) The lipophilic guanosine compound capable of chiral discrimination between ( <i>D</i> ) and ( <i>L</i> ) <i>N</i> -(2,4-dinitrophenyl)-tryptophan, and b) potassium salt of <i>N</i> -(2,4-dinitrophenyl)-( <i>L, D</i> )-tryptophan. ....	24
<b>Figure 1.23.</b>	Preparation of a library of acylhydrazones (E, F) and G-quartet hydrazones (G, H) from the reaction of guanosine hydrazide A and serine hydrazide B with aldehydes C and D. ....	25
<b>Figure 1.24.</b>	A schematic representation of the reaction of acridone (A), peptide (P), and glutathione (G, both oxidized and reduced forms) in the presence and absence of the G-quadruplex template.....	27
<b>Figure 1.25.</b>	Synthesis of a unimolecular G-quadruplex using reversible olefin metathesis. ...	28
<b>Figure 1.26.</b>	Interconversion of a bisguanidine compound between sol and gel states.....	29
<b>Figure 1.27.</b>	A) polydiacetylene (PDA) liposome, B) a polydiacetylene (PDA) liposome functionalized with G-rich oligonucleotides (non fluorescent) and the corresponding microarray, and C) activated liposome (red fluorescent) due to the formation of G-quadruplex. ....	31
<b>Figure 1.28.</b>	A schematic representation of the on/off nanoswitch used for the electrochemical detection of potassium ion. ....	31
<b>Figure 1.29.</b>	The structure of the thrombin-binding aptamer containing two pyrene moieties and the resulting chair-type conformation. The interaction of the fluorophores is only possible in the presence of potassium ions. ....	32
<b>Figure 1.30.</b>	Schematic illustration of the three-component duplex. Addition of thrombin leads to the formation of the G-quadruplex, which coincides with the release of the Q-strand and an enhancement in the fluorescence.....	33
<b>Figure 1.31.</b>	A) Formation of monolayer, B) addition and capture of thrombin, C) addition and binding of the secondary aptamer, and D) oxidation and detection of Cd <sup>2+</sup> with an ion-selective electrode.....	34
<b>Figure 1.32.</b>	Proposed mechanisms for the recognition of the G-quadruplex-based molecular beacon. ....	36
<b>Figure 1.33.</b>	A schematic representation of the deoxyribozyme catalyzing the photocleavage of a thymine cyclobutane dimer. The G-quadruplex absorbs and transmits the low-energy radiation (~305 nm) which is outside the normal absorption of DNA (>250 nm).....	37
<b>Figure 1.34.</b>	A proline modified G-quadruplex catalyzes the aldol reaction between acetone and an aldehyde-appended porphyrin. ....	38
<b>Figure 1.35.</b>	Hybrid tetramolecular G-quadruplexes formed from DNA (blue) and PNA (Red). Two structures of diagonally opposite (A) or adjacent (B) are possible. ....	39
<b>Figure 1.36.</b>	a) Schematic representation of the bunch-oligonucleotides (Arrows indicate 5' → 3' polarity), b) G-quadruplexes obtained from bunch-oligonucleotides in different orientations. Structure <b>3b</b> was not formed and <b>4b</b> and <b>4c</b> were obtained in minor quantities. ....	41

<b>Figure 2.1.</b>	The important 2D NMR correlations observed for compound <b>10c</b> (shown broken apart for the convenience of illustrating the correlations). ....	57
<b>Figure 2.2.</b>	400 MHz $^1\text{H}$ - $^1\text{H}$ COSY of <b>10c</b> at 25 °C in DMSO- $d_6$ . ....	60
<b>Figure 2.3.</b>	400 MHz $^1\text{H}$ - $^{13}\text{C}$ HMQC of <b>10c</b> at 25 °C in DMSO- $d_6$ . The signals corresponding to diastereotopic pairs of H5'a/ H5'b, H <sub>a</sub> /H <sub>b</sub> , and H <sub>in</sub> /H <sub>out</sub> are circled. ....	61
<b>Figure 2.4.</b>	400 MHz $^1\text{H}$ - $^1\text{H}$ COSY of <b>10c</b> at 25 °C in CDCl <sub>3</sub> . ....	62
<b>Figure 2.5.</b>	400 MHz $^1\text{H}$ - $^{13}\text{C}$ HMQC of <b>10c</b> at 25 °C in CDCl <sub>3</sub> . The signals corresponding to diastereotopic pairs of H5'a/ H5'b, H <sub>a</sub> /H <sub>b</sub> , and H <sub>in</sub> /H <sub>out</sub> are circled. ....	63
<b>Figure 2.6.</b>	a) NOEs expected for 1,4 and 1,5 triazoles. b) NOEs observed for a solution of <b>10a</b> in DMSO at 400 MHz at 25 °C. ....	64
<b>Figure 2.7.</b>	400 MHz $^1\text{H}$ NMR of <b>10c</b> at 25 °C in a) CDCl <sub>3</sub> , and b) DMSO- $d_6$ . ....	65
<b>Figure 2.8.</b>	400 MHz $^1\text{H}$ NMR of <b>10c</b> at -40 °C in CDCl <sub>3</sub> indicating the H-bonded (NH <sub>2b</sub> ), and the non-H-bonded (NH <sub>2a</sub> ) amino signals. ....	66
<b>Figure 2.9.</b>	a) Inter and intra-base NOE correlations in a G-quartet, b) Intra-base NOE correlations in <i>syn</i> (strong H1'/H8 and weak H2'/H8) and <i>anti</i> (medium H1'/H8 and strong H2'/H8) conformers, <sup>24</sup> c) NOEs indicative of the formation of G-quartet, and d) NOEs indicative of the <i>syn</i> conformation at 400 MHz in CDCl <sub>3</sub> at -40 °C. ....	67
<b>Figure 2.10.</b>	Variable-temperature experiments on a solution of TASQ <b>10c</b> in CDCl <sub>3</sub> at 400 MHz. Arrows mark signals of: red: water, black: H-bonded NH <sub>2</sub> , and yellow: the average NH <sub>2</sub> . ....	69
<b>Figure 2.11.</b>	Changes in the chemical shift of coalesced NH <sub>2</sub> signal of <b>10c</b> in CDCl <sub>2</sub> -CDCl <sub>2</sub> from 50-100°C. ....	70
<b>Figure 2.12.</b>	The donor and acceptor faces of guanine in a G-quartet unit. ....	71
<b>Figure 2.13.</b>	The amino (NH <sub>2</sub> ) and imino (NH) stretching region (3200-3600 cm <sup>-1</sup> ) of the infrared spectrum of a 3×10 <sup>-2</sup> M solution of TASQ <b>10c</b> in CDCl <sub>3</sub> at RT. The association bands at 3310 and 3470 cm <sup>-1</sup> are due to H-bonding. ....	73
<b>Figure 2.14.</b>	CD spectra of a 0.2 mM solution of TASQ <b>10c</b> in chloroform and DMSO. ....	75
<b>Figure 2.15.</b>	The translational diffusion of a molecule from point 1 to point 2, which is independent of its path. ....	76
<b>Figure 2.16.</b>	a) A simple gradient pulse sequence, b) The BPLED sequence. In this sequence, short gradients of opposite polarity are used, which are separated by a 180° pulse. This combination reduces the eddy current effects (induced by gradient pulses), and improves the line shape. ....	78
<b>Figure 2.17.</b>	Normalized signal decay as a function of the b value at 295 K for a) <b>10c-16c</b> in DMSO- $d_6$ , b) <b>10c-16c</b> in CDCl <sub>3</sub> . The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value = (2·π·γ·g·δ) <sup>2</sup> (Δ-δ/3) s/m <sup>2</sup> ]. ....	82
<b>Figure 2.18.</b>	Normalized signal decay as a function of the g (gradient strength) at 295 K for a) the H1' signal of <b>10c</b> in DMSO- $d_6$ , and b) the H1' signal of <b>10c</b> in CDCl <sub>3</sub> . ....	83
<b>Figure 3.1.</b>	$^1\text{H}$ NMR spectra of TASQ-cation assemblies in CDCl <sub>3</sub> at 400 MHz at 25 °C. The region between 6 and 12 ppm's has been assigned to make comparison easier. The signals labeled "p" and "s" indicate picrate and solvent respectively. TASQ <b>10c</b> ·Cs <sup>+</sup> gave two sets of signals, one of which has been marked with asterisks. ....	101
<b>Figure 3.2.</b>	$^1\text{H}$ - $^1\text{H}$ COSY of TASQ <b>10c</b> ·Na <sup>+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C. ....	103
<b>Figure 3.3.</b>	TASQ <b>10c</b> ·Na <sup>+</sup> shows NOEs similar to those shown by the cation-free system. These effects are indicative of a) the formation of G-quartet and, b) the <i>syn</i> conformation at 400 MHz in CDCl <sub>3</sub> at -40 °C. The NH <sub>2</sub> signals which appear broad at RT, sharpen up and become clearly visible at low temperature. ....	104

<b>Figure 3.4.</b>	Picrate signal integration for TASQ <b>10c</b> ·Na <sup>+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C....	105
<b>Figure 3.5.</b>	Normalized signal decay as a function of the b value at 295 K for <b>10c</b> ·Na <sup>+</sup> - <b>16c</b> in CDCl <sub>3</sub> . The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value = $(2 \cdot \pi \cdot \gamma \cdot g \cdot \delta)^2 (\Delta \cdot \delta / 3) \text{ s/m}^2$ ] .....	106
<b>Figure 3.6.</b>	Normalized signal decay as a function of the gradient strength (g) at 295 K for a) the H1' signal of <b>10c</b> ·Na <sup>+</sup> , and b) the H1' signal of <b>16c</b> in CDCl <sub>3</sub> . .....	106
<b>Figure 3.7.</b>	<sup>1</sup> H- <sup>1</sup> H COSY of TASQ <b>10c</b> ·Sr <sup>2+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C.....	108
<b>Figure 3.8.</b>	Portions of a 100 ms NOESY spectrum of TASQ <b>10c</b> ·Sr <sup>2+</sup> acquired at 400 MHz at -40 °C in CDCl <sub>3</sub> indicative of a) the formation of G-quartet, and b) the <i>syn</i> conformation. The NH <sub>2</sub> signals which are broadened up at RT are visible at low temperature. ....	109
<b>Figure 3.9.</b>	Picrate signal integration for TASQ <b>10c</b> ·Sr <sup>2+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C....	109
<b>Figure 3.10.</b>	Normalized signal decay as a function of the b value at 295 K for <b>10c</b> ·Sr <sup>2+</sup> - <b>16c</b> in CDCl <sub>3</sub> . The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value = $(2 \cdot \pi \cdot \gamma \cdot g \cdot \delta)^2 (\Delta \cdot \delta / 3) \text{ s/m}^2$ ] .....	110
<b>Figure 3.11.</b>	Normalized signal decay as a function of the gradient strength (g) at 295 K for a) the H1' signal of <b>10c</b> ·Sr <sup>2+</sup> , and b) the H1' signal of <b>16c</b> in CDCl <sub>3</sub> .....	111
<b>Figure 3.12.</b>	Partial <sup>1</sup> H NMR spectrum of TASQ <b>10c</b> ·K <sup>+</sup> at 400 MHz at -40 °C displaying three sets of signals. The two sets of small signals related to a second “new” species has been marked with asterisks. The signals labeled “p” and “s” corresponds to the picrate and the solvent, respectively.....	112
<b>Figure 3.13.</b>	Portions of a 100 ms NOESY spectrum of TASQ <b>10c</b> ·K <sup>+</sup> acquired at 400 MHz in CDCl <sub>3</sub> .....	112
<b>Figure 3.14.</b>	Picrate signal integration for TASQ <b>10c</b> ·K <sup>+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C.....	113
<b>Figure 3.15.</b>	Partial <sup>1</sup> H- <sup>1</sup> H NOESY of TASQ <b>10a</b> ·Cs <sup>+</sup> acquired at 400 MHz at -40 °C in CDCl <sub>3</sub> . Signals shown are indicative of the formation of G-quartet. ....	114
<b>Figure 3.16.</b>	Picrate signal integration for TASQ <b>10c</b> ·Cs <sup>+</sup> at 400 MHz in CDCl <sub>3</sub> at 25 °C. ....	115
<b>Figure 3.17.</b>	Normalized signal decay as a function of the b value at 295 K for <b>10c</b> ·Cs <sup>+</sup> - <b>10c</b> in CDCl <sub>3</sub> . The diffusion coefficients (D) were calculated for the representative signal of NH. [b value = $(2 \cdot \pi \cdot \gamma \cdot g \cdot \delta)^2 (\Delta \cdot \delta / 3) \text{ s/m}^2$ ] .....	117
<b>Figure 3.18.</b>	Normalized signal decay as a function of the gradient strength (g) at 295 K for a) the NH signal of <b>10c</b> ·Cs <sup>+</sup> , and b) the NH signal of <b>10c</b> in CDCl <sub>3</sub> . ....	118
<b>Figure 3.19.</b>	Proposed G-quartet-based structures of TASQ·cation assemblies a) isolated G-quartet, b,d) homodimers, and c) heterodimer. Only structures consistent with (a) and (c) were observed.....	119
<b>Figure 3.20.</b>	ESI-MS spectrum of TASQ <b>10a</b> ·Sr <sup>2+</sup> in positive mode. Two species of [TASQ <b>10a</b> ·Sr] <sup>2+</sup> (m/z = 1145) and [(TASQ <b>10a</b> ) <sub>2</sub> ·Sr] <sup>2+</sup> (m/z = 2246) were observed in the gas-phase. ....	121
<b>Figure 3.21.</b>	MALDI-MS spectrum of TASQ <b>10a</b> ·Sr <sup>2+</sup> in positive mode. Two species of [TASQ <b>10a</b> ·Sr] <sup>+</sup> (m/z = 2290) and [(TASQ <b>10a</b> ) <sub>2</sub> ·Sr] <sup>+</sup> (m/z = 4492) were observed in the gas-phase (note that MALDI typically generates only singly charged species) .....	121
<b>Figure 3.22.</b>	a) <sup>1</sup> H NMR spectrum of pure TASQ <b>10a</b> ·Cs <sup>+</sup> , b) <sup>1</sup> H NMR spectrum of pure TASQ <b>10a</b> ·Na <sup>+</sup> , c) <sup>1</sup> H NMR spectrum recorded after extraction of TASQ <b>10a</b> with an equimolar mixture of sodium and cesium picrates. ....	124
<b>Figure 3.23.</b>	Schematic representation of the conversion of TASQ <b>10c</b> ·Na <sup>+</sup> to TASQ <b>10c</b> upon addition of [2.2.2] cryptand. ....	125

<b>Figure 3.24.</b>	a) $^1\text{H}$ NMR spectrum of a 2.4 mM solution of TASQ <b>10c</b> ·Na <sup>+</sup> at 400 MHz in CDCl <sub>3</sub> at 25°C, b) The spectrum recorded after addition of the 8-fold excess of [2.2.2] cryptand to the same solution (note the free picrate signal at ~8.8 ppm), c) $^1\text{H}$ NMR spectrum of a 2.4 mM of TASQ <b>10c</b> after addition of the 8-fold excess of [2.2.2] cryptand, and d) $^1\text{H}$ NMR spectrum of a solution of TASQ <b>10c</b> at 400 MHz in CDCl <sub>3</sub> at 25°C.....	126
<b>Figure 3.25.</b>	CD spectra of a 0.2 mM solution of TASQ <b>10c</b> and TASQ <b>10c</b> -cation assemblies in chloroform. ....	128
<b>Figure 4.1.</b>	Tris-(benzyltriazolylmethyl)amine (TBTA) .....	145
<b>Figure 4.2.</b>	The upper and lower rims of the cavitand.....	149
<b>Figure 4.3.</b>	HPLC chromatograms of the crude (top), and the pure <b>20a</b> (bottom).....	168
<b>Figure 4.4.</b>	ESI-MS spectrum of <b>20a</b> in negative mode. The signals at ( $m/z = 1278.2$ , 1300.1, and 1316.1) correspond to [M-H] <sup>-</sup> , [M+Na-2H] <sup>-</sup> , and [M+K-2H] <sup>-</sup> respectively. ....	168
<b>Figure 4.5.</b>	HPLC chromatograms of the crude (top), and pure <b>20b</b> (bottom). ....	169
<b>Figure 4.6.</b>	MALDI-MS spectrum of <b>20b</b> in positive mode. The signal at ( $m/z = 1888.8$ ) corresponds to [M+H] <sup>+</sup> . ....	169
<b>Figure 4.7.</b>	ESI-MS spectrum of <b>34</b> in positive mode. Only one species was observed in the solution. The signals at ( $m/z = 1457.7$ ) and ( $m/z = 1558.9$ ) correspond to [M+H] <sup>+</sup> and [M+NEt <sub>3</sub> +H] <sup>+</sup> respectively.....	171
<b>Figure 4.8.</b>	The 200 MHz $^{31}\text{P}$ NMR spectrum of a solution of <b>5a</b> reacted with 2-cyanoethyl- <i>N,N</i> -diisopropylchloro phosphoramidite in CD <sub>2</sub> Cl <sub>2</sub> at 25 °C. The signal at ~149.9 ppm corresponds to the product. (Note that the P-diastereomers are not resolved at 200 MHz).....	171
<b>Figure 5.1.</b>	a) A ditopic TASQ monomer, b) Three possible arrays derived from self-assembly of compound <b>39</b> : a linear polymer <b>39a</b> , a sandwiched dimer <b>39b</b> , and a stacked polymer <b>39c</b> (only one stack is shown).....	181

## List of Schemes

<b>Scheme 2.1.</b>	Thermal vs. catalytic 1,3-dipolar cycloaddition reaction.....	53
<b>Scheme 2.2.</b>	The synthesis of cavitands <b>6a-c</b> . ....	54
<b>Scheme 2.3.</b>	The synthesis of 5'-azido guanosine <b>9</b> . ....	55
<b>Scheme 2.4.</b>	The synthesis of TASQs <b>10a-c</b> . ....	56
<b>Scheme 2.5.</b>	The Synthesis of Compound <b>16c</b> . ....	80
<b>Scheme 2.6.</b>	The Synthesis of Nucleoside <b>15</b> . ....	91
<b>Scheme 4.1.</b>	Proposed synthesis of an all parallel hydrophilic TASQ with the sequences of 5'-TG <sub>4</sub> T-3' and 5'-G <sub>4</sub> -3' using the oligonucleotide approach.....	140
<b>Scheme 4.2.</b>	The solid-phase functionalization of <b>17a-b</b> . ....	142
<b>Scheme 4.3.</b>	Proposed synthesis of a hydrophilic TASQ using the water-soluble cavitand approach. ....	148
<b>Scheme 4.4.</b>	The synthesis of phosphate-footed tetrakis ( <i>O</i> -propargyl) cavitand <b>30</b> . ....	150
<b>Scheme 4.5.</b>	Proposed synthesis of a hydrophilic TASQ using the phosphoramidite approach. ....	152
<b>Scheme 4.6.</b>	The transient protection of <b>7</b> . ....	153
<b>Scheme 4.7.</b>	The synthesis of tetrakisphosphoramidite cavitand <b>34</b> . ....	153
<b>Scheme 4.8.</b>	The synthesis of <i>N</i> <sup>2</sup> -isobutyryl-2',3'- <i>O</i> -isopropylidene guanosine phosphoramidite.....	154
<b>Scheme 4.9.</b>	Proposed synthesis of a hydrophilic TASQ using the phosphoramidite approach. ....	155
<b>Scheme 5.1.</b>	a) Proposed synthesis of a pentameric TASQ analog using the “click” reaction, b) Isoguanosine pentamer. ....	182
<b>Scheme 5.2.</b>	Proposed reactants for the synthesis of a hydrophilic TASQ using the “click” reaction and their preparation methods.....	185

## List of Abbreviations

AFM	atomic force microscopy
aq.	aqueous
Ar	aromatic
BBI	broadband inverse (probe)
BPLED	bipolar pulse longitudinal eddy current delay
Bn	benzyl
Bu	butyl
Bz	benzyl
CD	circular dichroism
COSY	correlation (NMR) spectroscopy
CPG	controlled pore glass
CPK	Corey-Pauling-Koltun
CTAB	cetyltrimethylammonium bromide
$\Delta\delta$	change in chemical shift
d	doublet
DCC	dynamic covalent chemistry
DCL	dynamic covalent library
DCM	dichloromethane
DCTB	2-[(2E)-3-(4-tert-butylphenyl)-2-methylprop-2-enylidene]malononitrile
DHB	2,5-dihydroxybenzoic acid
DMA	<i>N,N</i> -dimethylacetamide
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulfoxide
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
EDTA	ethylenediamine
ESI	electrospray ionization
Et	ethyl (-CH <sub>2</sub> CH <sub>3</sub> )
EtOAc	ethyl acetate
EtOH	ethanol
eq.	equivalents
EXAFS	extended X-ray absorption fine structure
FRET	Förster resonance energy transfer
GMP	guanosine monophosphate
HIV	human immunodeficiency virus
HMQC	heteronuclear multiple quantum coherence
HPA	hydroxylpicolinic acid
HPLC	high performance (pressure) liquid chromatography
i-Pr	isopropyl
IR	infrared
ISE	ion-selective electrode
L	litre
m	multiplet
M	parent mass (mass spectra) or moles per litre

MALDI	matrix assisted laser desorption ionization
MB	molecular beacon
Me	methyl (-CH <sub>3</sub> )
MeCN	acetonitrile
MeOH	methanol
MHz	megahertz
Ms	methanesulfonyl
MS	mass spectrometry or spectrum
MW	molecular weight
<i>n</i> -BuLi	<i>n</i> -butyllithium
NBS	<i>N</i> -bromosuccinimide
nm	nanometer(s)
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhancement spectroscopy
PDA	polydiacetylene
PFG	pulse field-gradient
Ph	phenyl group
PNA	peptide-nucleic acid
q	quartet ( <sup>1</sup> H NMR)
RF	radio frequency
RNA	ribonucleic acid
RP-HPLC	reverse-phase high performance (pressure) liquid chromatography
RT	room temperature
s	singlet ( <sup>1</sup> H NMR)
SANS	small-angle neutron scattering
t	triplet
TASP	template assembled synthetic protein
TASQ	template assembled synthetic G-quartet
TBA	thrombin-binding aptamer
TBAF	tetrabutylammonium fluoride
TBDMS	tert-butyltrimethylsilyl
TBTA	tris-(benzyltriazolylmethyl)amine
t-Bu	tert-butyl
TCEP	tris(2-carboxyethyl)phosphine hydrochloride
TEA	triethylamine
TFA	2,2,2-trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMPyP4	tetramethylpyridinium porphyrin
TOF	time of flight
Ts	para-toluenesulfonyl
UV	ultraviolet
vs.	versus
v/v	volume per volume

## **Acknowledgements**

First, I would like to thank my supervisor, Professor John Sherman, who shaped my mind as a researcher and taught me to think critically and not to be satisfied with the minimum. I really appreciate his patience, constructive discussions and the time he spent reading and improving the manuscripts. I am also grateful to Dr. David Perrin for inspiring this project, and his helpful suggestions and constant encouragement.

Also, I wish to thank both present and past members of the Sherman group for creating a friendly atmosphere in the lab. Special thanks to Alfredo Franco Cea, Jon Freeman and Grant Bare for proofreading sections of this thesis.

I would like as well to acknowledge the staff of the NMR laboratory, Dr. Nick Burlinson, Dr. Maria Ezhova and Zorana Danilovic for providing assistance and sharing their valuable experience, Bert Meuller for conducting atomic absorption experiments, Fred Rosell for helping with his CD apparatus and the staff of the mass spectrometry laboratory and biological services.

I want to thank Dr Marco Ciufolini as well for his helpful suggestions regarding the synthesis of the water-soluble cavitand.

Finally, I would also like to express my sincere thanks, appreciation and respect to my family and friends.

## **Co-Authorship**

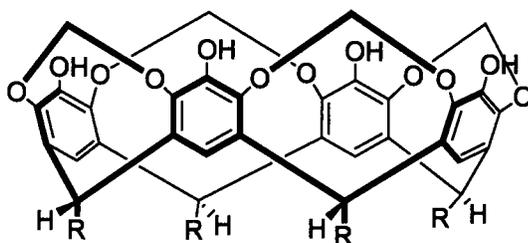
This thesis work has been conducted and written solely by the author and not in collaboration with other parties (including experimental design and data analyses). The published and submitted manuscripts have been prepared by the author and have been edited by Professor John Sherman.

## CHAPTER 1: Introduction

### 1.1 Background

Supramolecular chemistry deals with non-covalent interactions between molecules. One of its challenges is the organization of small molecules into well-defined assemblies.<sup>1</sup> One way to approach this problem is to use a molecular template to which structural units can be attached.<sup>2</sup> The molecules which are generally used as templates are rigid molecules that direct the building blocks in well-defined spatial arrangements. Numerous supramolecular systems containing both rigid and flexible subunits have been designed and synthesized following this approach.<sup>3</sup> This strategy has also been used to address the protein folding problem.<sup>4</sup> The cavitands, rigid macrocycles with an enforced cavity, are ideal scaffolds which have been used for synthesis of *de novo* proteins (Figure 1.1). The resulting four-helix bundles called caviteins, have been studied in our group and have shown a high degree of stability compared to other template-assembled synthetic proteins (TASPs).<sup>5,6</sup>

**Figure 1.1.** A tetrol cavitand



## 1.2 Thesis Goals

In the present study we are utilizing the concept of template-assembly along with H-bonding interactions towards the directed organization of G-quartet assemblies. G-quartets are cyclic tetramers of guanines which are important in biology and supramolecular chemistry.<sup>7</sup> The major goals of this project are the followings:

1) To design and synthesize a family of cavitand-linked guanines such that the cavitand helps direct the guanines into a G-quartet. We have named these compounds template-assembled synthetic G-quartets (TASQs).

2) To demonstrate the utility of the cavitand template in controlling the morphology of G-quadruplexes (stacks of G-quartets).

## 1.3 Thesis Overview

This thesis is organized in five chapters. Chapter 1 will describe the supramolecular and biological significance of the G-quartets and G-quadruplexes, and will present selected examples of the relevant works. Chapter 2 will detail the design, synthesis, and solution properties of the first lipophilic template-assembled synthetic G-quartets (TASQs) which are capable of forming a cation-free G-quartet in solution. Chapter 3 will discuss the cation-complexation behavior of the TASQs, and investigate the preparation and characterization of the TASQ·cation assemblies. Chapter 4 will present the attempts at synthesizing a hydrophilic TASQ including the synthesis of a water-soluble phosphate-footed tetrapropargyl ether cavitand. Chapter 5 will present the thesis summary, and will suggest

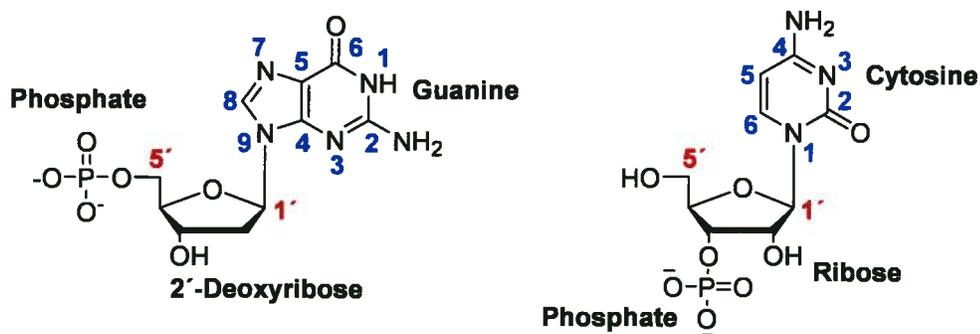
directions for future research. An introduction on nucleotides will be presented below, followed by a discussion on the structure of G-quartets.

## 1.4 G-Quartet in Supramolecular Chemistry and Biology

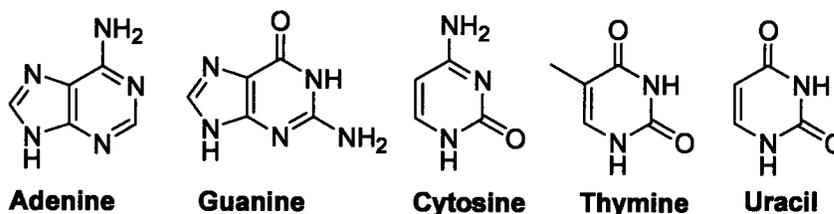
### 1.4.1 Properties of Nucleotides

Nature has always been a source of inspiration for supramolecular chemists. Vital biopolymers such as DNA and RNA are both constructed from smaller units selected by the powerful mechanism of evolution. These smaller units, called nucleotides, (Figure 1.2) are composed of a heterocyclic base (purine or pyrimidine), a sugar (*D*-ribose or *D*-2'-deoxyribose), and a phosphate group (at 3' or 5' positions). Figure 1.3 shows the nucleobases of DNA and RNA which are all capable of forming highly ordered H-bonds.<sup>8,9</sup>

**Figure 1.2.** Two sample nucleotides of guanine and cytosine. The numbering schemes for the base and sugar subunits are also presented.

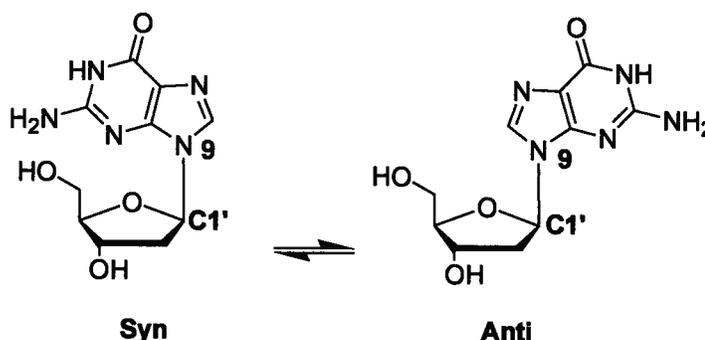


**Figure 1.3.** The purine (adenine, guanine) and pyrimidine (cytosine, thymine, uracil) bases of nucleic acids.



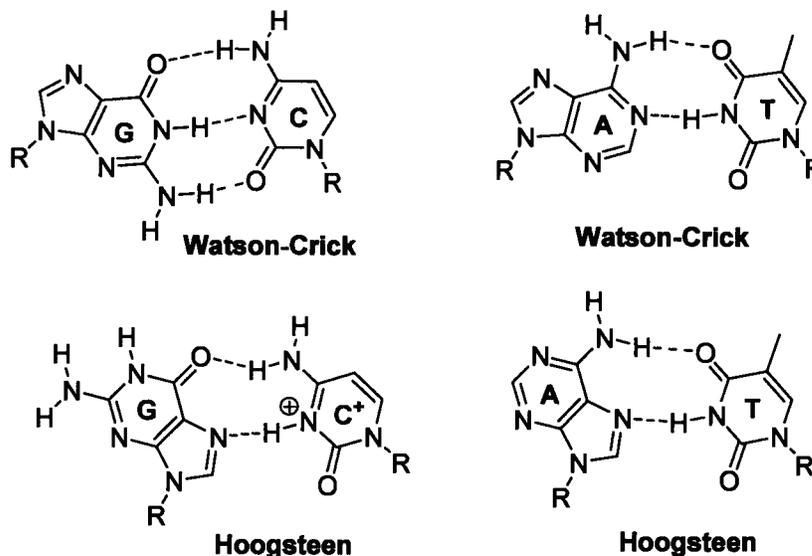
The nucleobases in DNA and RNA are all connected to the ribose moieties through an *N*-glycosidic bond and can exist in two different conformations known as *syn* and *anti* (Figure 1.4). The *anti* conformation is the dominant conformation in the naturally occurring DNA.<sup>8,9</sup>

**Figure 1.4.** The *syn* and *anti* conformations of guanosine.



Because of their geometry, all bases in DNA can form Watson-Crick H-bonds in which a purine pairs with a pyrimidine base (Figure 1.5). There are also other modes of H-bonding including the Hoogsteen pairing, which is probably more significant than the other non-Watson-Crick H-bonds (Figure 1.5).<sup>10</sup> This mode of H-bonding plays a key role in the formation of the G-quartet.<sup>11</sup>

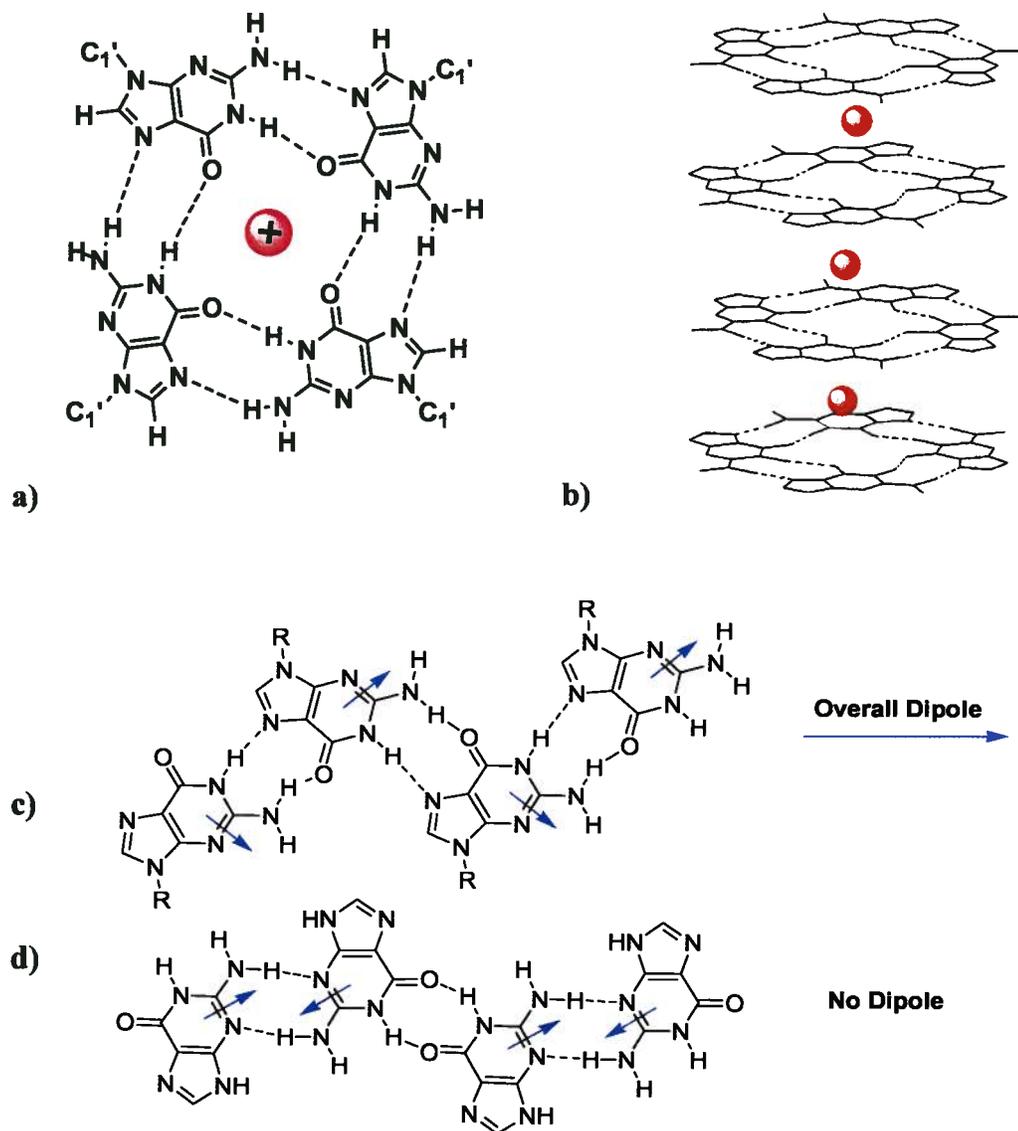
**Figure 1.5.** Watson-Crick and Hoogsteen base-pairing of nucleobases. Cytosine has to be protonated to be able to participate in Hoogsteen pairing with guanine or another molecule of cytosine. This mode of H-bonding is important in the formation of triplexes and i-motifs.<sup>12</sup>



#### 1.4.2 Guanine Self-Assembly: G-Quartet vs. G-Ribbon

Guanine bases, present in guanosine compounds, guanine-rich strands of DNA and RNA, and DNA-like polymers, are able to self-associate and form a tetrameric structure called a G-quartet (Figure 1.6a).<sup>7,11,13</sup> Stacks of G-quartets form a three dimensional structure known as a G-quadruplex (Figure 1.6b).<sup>14,15</sup> As shown in Figure 1.6a, each G-quartet is stabilized by eight hydrogen bonds. Some positively charged metal cations such as Na<sup>+</sup>, K<sup>+</sup>, and Sr<sup>2+</sup> can also be sandwiched between the planes of the G-quartets and their inclusion increases the stability of the system.<sup>16,17</sup> These cations are coordinated by the carbonyl oxygen atoms of the G-quartets. Guanine aggregation in the absence of cations (in lipophilic solvents or on the solid phases) generally results in the formation of ribbon-like structures.<sup>18</sup> Depending on the solvent and the sugar substituents, one of the structures of (c) or (d) might be formed (Figures 1.6c and 1.6d).

**Figure 1.6.** A schematic representation of a) G-quartet, b) G-quadruplex (the G-quartet units are rotated about 30° in respect to each other). Images (c) and (d) depict G-ribbons (arrows indicate dipole moments).



### 1.4.3 Nucleic Acid G-Quadruplexes

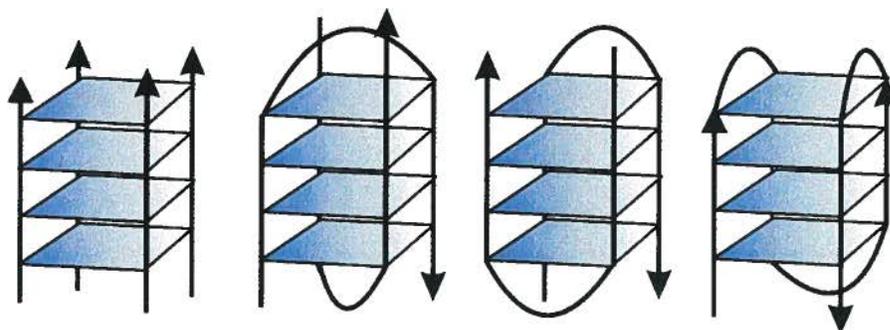
G-quartet formation has been known for many years, but little attention was paid to this phenomenon.<sup>19</sup> The early studies on G-quartets date back to 1962 when guanine tetramer was

discovered as the main component of the gels formed from the 3' and 5'-guanosinemonophosphate (GMP) in water.<sup>11</sup> There is currently considerable interest in this area of research because of the potential role of these structures in many biological events.<sup>7</sup> G-rich sequences are abundant in some biologically important regions, such as the telomeric ends of eukaryotic chromosomes, the promoter regions of DNA,<sup>20,21</sup> and the immunoglobulin switch regions<sup>22</sup>, and are capable of forming G-quadruplexes *in vitro*. There is also evidence for their existence and function *in vivo*.<sup>23</sup>

Telomeric DNA shrinks with each successive division of cells, thus limiting the number of divisions a cell can undergo. In cancerous cells, the telomerase enzyme elongates the telomeric DNA, enabling more divisions of cells. This enzyme is only active in cancerous cells but inactive in most somatic cells. It has been suggested that formation of a G-quadruplex could inhibit cancer because the telomerase enzyme only binds to the unfolded telomere. When a single stranded telomere folds over on itself and forms a G-quadruplex, the action of the telomerase is terminated. Therefore, ligands that can selectively stabilize or bind to G-quadruplexes are of utmost importance. Some of these ligands have been shown to accelerate the folding of a single strand of DNA to a G-quadruplex.<sup>24,25</sup>

G-quadruplex DNA forms a wide array of structures and can be divided into a number of subclasses. These topological structures are important because different G-quadruplex topologies may differ in their biological roles. The diversity of quadruplex structures is attributed to the variation in strand stoichiometry (one, two or four strands), strand polarity or direction (parallel vs. antiparallel), glycosidic conformations (*syn* and *anti*), and lastly connecting loops and cations (Figure 1.7).<sup>7,26</sup>

**Figure 1.7.** a) Tetramolecular, b,c) bimolecular and d) unimolecular G-quadruplexes (arrows indicate 3'→5' polarity).



#### 1.4.4 Lipophilic G-Quadruplexes

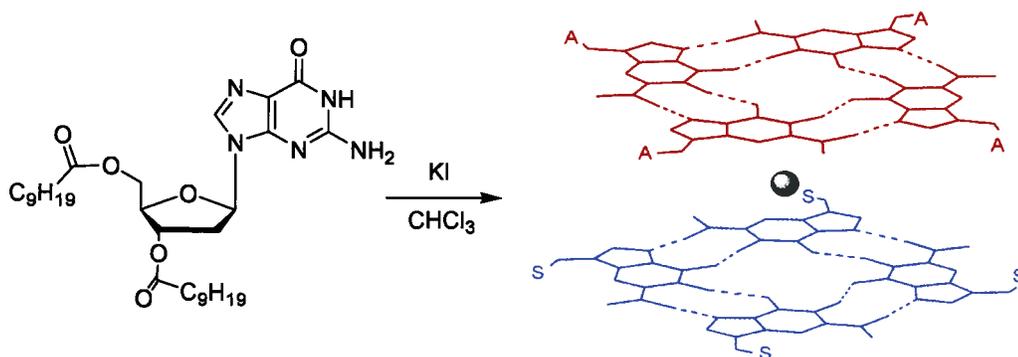
Compared to the studies conducted on nucleic acid G-quadruplexes, the research on lipophilic systems has been slow for many years, mainly due to solubility problems. With the introduction of the solubilizing protecting groups such as silyl ethers, the research in this area has been given a new boost.<sup>13</sup> Gottarelli's, Davis's and Spada's labs have been particularly active in this area of research and contributed extensively to the development of this field. Lipophilic G-quadruplexes are useful materials for constructing supramolecular systems and are also ideal models for the nucleic acid G-quadruplexes.<sup>13,27</sup>

In 1995, Gottarelli and colleagues demonstrated that the lipophilic guanosine compounds are capable of extracting alkali metal picrates from aqueous phase to chloroform.<sup>28</sup> Later on they established the structure of the G-ribbons (shown in Figures 1.6c and 1.6d)<sup>18</sup> and in collaboration with Davis's lab solved the structure of the first lipophilic G-quadruplex in solution.<sup>29</sup> We will start this section with a review of the important findings from the above mentioned research groups followed by a brief discussion on the role of cations.

### 1.4.4.1 Structure of Lipophilic G-Quadruplexes

One of the early studies conducted on 3',5'-didecanoyl-2'-deoxyguanosine in 1999 showed that this lipophilic guanosine compound forms an octamer in the presence of potassium ions (Figure 1.8).<sup>29</sup> The elimination of the phosphate groups ensures that the cations are only extracted by the guanine bases and coordinated to the central core of the G-quartet.

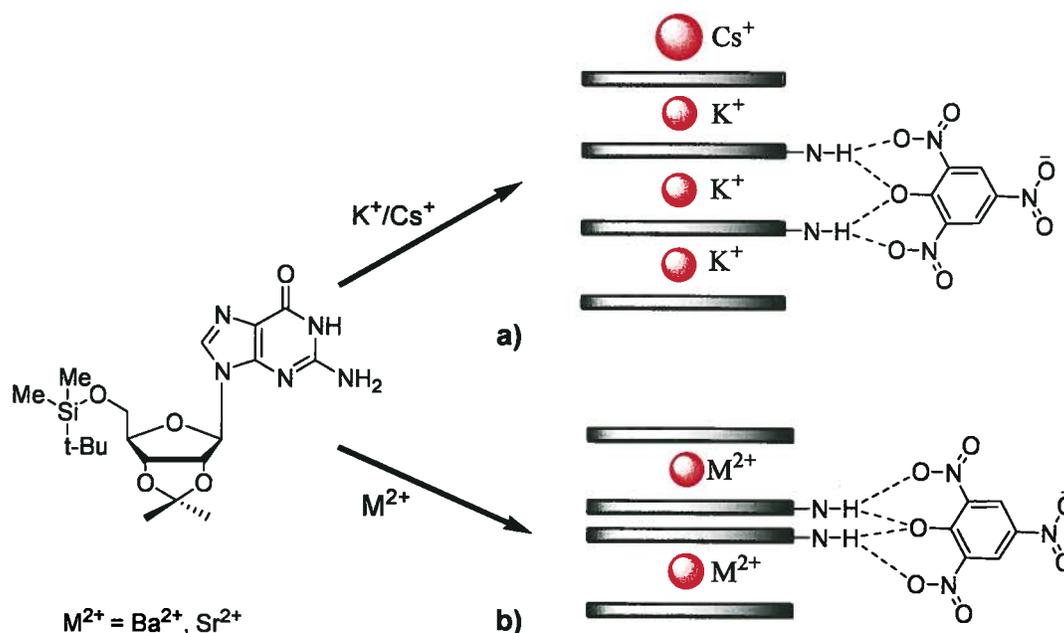
**Figure 1.8.** A schematic representation of an octamer formed from 3',5'-didecanoyl-2'-deoxyguanosine.



Gottarelli, Davis and colleagues demonstrated based on an extensive NMR analysis that the observed octamer is made of an all-*anti* and an all-*syn* tetramer sandwiching a central cation (Figure 1.8). The researchers also concluded that the formation of the octamer is the first stage in the formation of the higher order guanine aggregates. This conclusion was completely in line with the previous observations made in hydrophilic systems.<sup>17</sup> Another important achievement was accomplished in 2000 when the first crystal structure of lipophilic G-quadruplexes was obtained.<sup>30</sup> This crystal was a hexadecamer formed from a solution of 5'-silyl-2',3'-isopropylidene guanosine and  $K^+$  and  $Cs^+$  picrates in acetonitrile (Figure 1.9a). It was composed of four stacked G-quartets ( $\sim 3.3$  Å apart) resembling the crystals obtained from the parallel

nucleic acid G-quadruplexes.<sup>31</sup> Three potassium ions were coordinated along the central axis and a cesium ion was positioned on the upper quartet providing a cap for the system. Recently, it was shown that the hexadecamer is also the dominant species in the solution.<sup>32</sup>

**Figure 1.9.** A schematic representation of the crystal structure of a lipophilic guanosine hexadecamer formed in the presence of a)  $K^+$  and  $Cs^+$  picrates, b)  $M^{2+}$  picrates. The H-bonding of the picrate anion to the quartets is also presented.



The lipophilic G-quadruplexes are capable of binding both monovalent and divalent cations.<sup>7,13</sup> Besides the complexes of the monovalent cations, Davis's group was also successful in obtaining the crystal structures of the hexadecamers formed from some divalent cations such as  $Sr^{2+}$ ,<sup>33</sup> and  $Ba^{2+}$ .<sup>34</sup> In contrast to the first crystal, the divalent cations were only found between the upper and lower octamers, but no cation was coordinated to the inner G-quartets (Figure 1.9b). This observation has also been made in the nucleic acid G-quadruplexes and is due to the greater repulsive forces between the divalent cations.<sup>35</sup> The crystal structures also revealed that the anions contribute to the stability of the lipophilic G-quadruplexes by interacting

with the non-H-bonded amino proton of the guanines. Therefore, the anion acts as a clip for the quartets and holds the system together like an anionic belt (Figure 1.9). Anions such as 2,6-dinitrophenol, which makes stronger H-bonds with the quartets, provide better kinetic and thermodynamic stability for the system.<sup>36</sup> In addition to anions, cations also influence the kinetic and thermodynamic stability of the lipophilic G-quadruplexes.<sup>37</sup> Compared to the complexes formed by the monovalent cations such as  $\text{Na}^+$  and  $\text{K}^+$ , some divalent cations such as  $\text{Sr}^{2+}$  and  $\text{Ba}^{2+}$  provide a stronger ion-dipole interaction and as a result form more stable complexes.<sup>34</sup>

#### 1.4.5 The Role of Cations

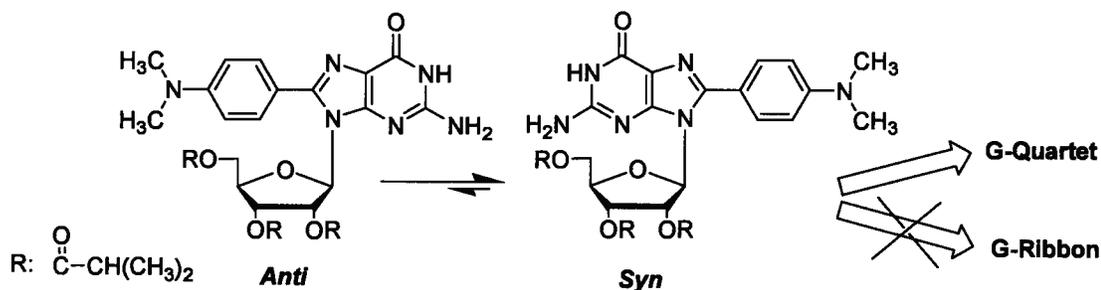
G-quadruplexes are templated and stabilized by cations. The coordination of cations to G-quartet layers was predicted and discovered about 30 years ago even before obtaining the crystal structures of the G-quadruplexes.<sup>17</sup> The early studies on 5'-GMP showed that in the presence of certain cations such as  $\text{K}^+$  and  $\text{Sr}^{2+}$ , a more stable gel with considerably higher melting temperature ( $T_m$ ) is obtained.<sup>38</sup> The following order of cation selectivity was proposed for 5'-GMP based on the optimum fit model:  $\text{K}^+ > \text{Na}^+, \text{Rb}^+ \gg \text{Li}^+, \text{Cs}^+$ .<sup>17</sup> Later studies on nucleic acid G-quadruplexes showed that, in addition to the size of the cations, the free energy of the coordination and, more importantly, the heat of dehydration of cations, contribute to the cation-selectivities.<sup>39,40</sup> The selectivities observed for nucleic acid G-quadruplexes are very similar to the one observed for 5'-GMP. A recent study on the telomeric G-quadruplex, for example, determined the following order:  $\text{Sr}^{2+} > \text{K}^+ > \text{Na}^+ \geq \text{Rb}^+ > \text{Li}^+ > \text{Cs}^+$ .<sup>41</sup> Cation-induced polymorphism of G-quadruplexes has also been under extensive investigation for years. Sen and Gilbert showed that a sodium-potassium switch can cause a structural transition from a parallel to an

antiparallel system.<sup>42</sup> Recent advances in the NMR techniques along with the crystallographic analyses and X-ray studies in solution-phase (EXAFS) have provided a wealth of information about the role of cations.<sup>31,35,43,44</sup> These studies are not limited only to hydrophilic G-quadruplexes. Wu and co-workers have successfully localized the <sup>23</sup>Na,<sup>45</sup> <sup>39</sup>K,<sup>46</sup> and <sup>87</sup>Rb<sup>47</sup> cations residing inside the lipophilic G-quadruplexes using solid-state NMR techniques. Aside from the direct detection of cations, the structural impact of cations on lipophilic G-quadruplexes has also been studied and will be discussed in the upcoming sections.

#### 1.4.6 Cation-Free G-Quartets

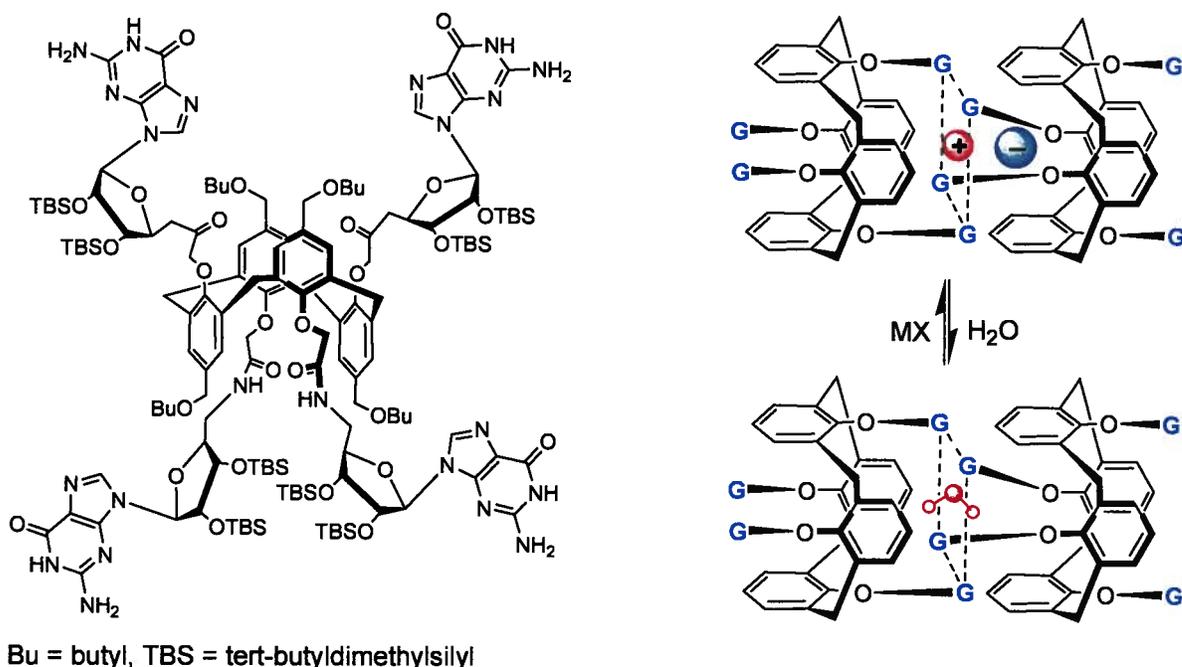
Although cations appear to be needed to stabilize G-quartet structures, this is not always the case. Sessler's group reported the formation of a cation-free G-quartet in the solution phase and solid-state from a guanosine compound substituted with a sterically demanding group on the C8 position (Figure 1.10).<sup>48</sup> The bulky substituent makes the formation of the *anti* conformation unfavorable and directs the formation of a cation-free G-quartet by blocking one of the faces of the guanine. The G-ribbons which have been reported to date have had *anti* conformations.

**Figure 1.10.** The *syn* conformation adopted by 8-(4-*N,N*-dimethylamino) guanosine directs the formation of a cation-free G-quartet.



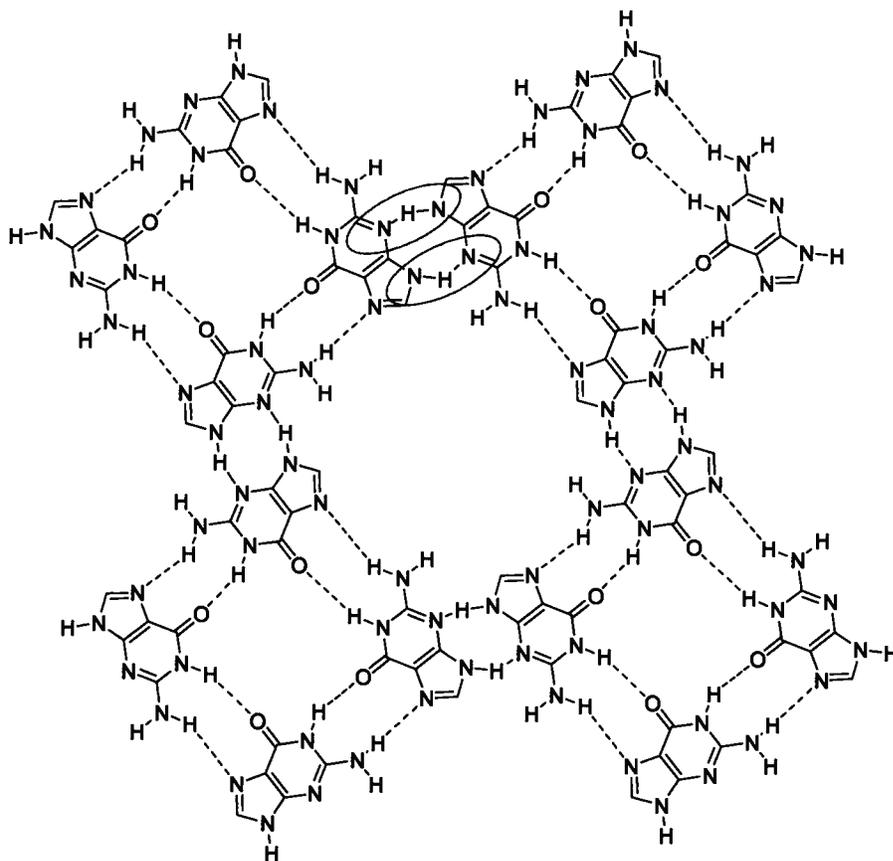
Recently, Davis et al. reported the synthesis and characterization of a calix[4]arene-guanosine conjugate.<sup>49</sup> The product is barely soluble in dry  $\text{CDCl}_3$ , but is more soluble in water-saturated  $\text{CDCl}_3$ . NMR analysis revealed that the conjugate forms a H-bonded dimer in this media which is held together by a water-bound G-quartet (Figure 1.11). The dimer was also found to be a ditopic receptor having two discrete binding sites for cations and anions.

**Figure 1.11.** a) Calix[4]arene-guanosine conjugate, and b) calix[4]arene-guanosine dimers.



Similar to guanosine, the guanine base itself has also been shown to form a cation-free G-quartet. More recently, Besenbacher and co-workers reported the formation of a network of cation-free G-quartets on a gold surface (Figure 1.12).<sup>50</sup> STM studies showed that the G-quartet was only the kinetic product and that the cyclic structure converted to G-ribbons upon annealing.

**Figure 1.12.** The network of cation-free G-quartet made from guanine on gold surfaces. The H-bonds that connect the adjacent G-quartets are marked with circles.

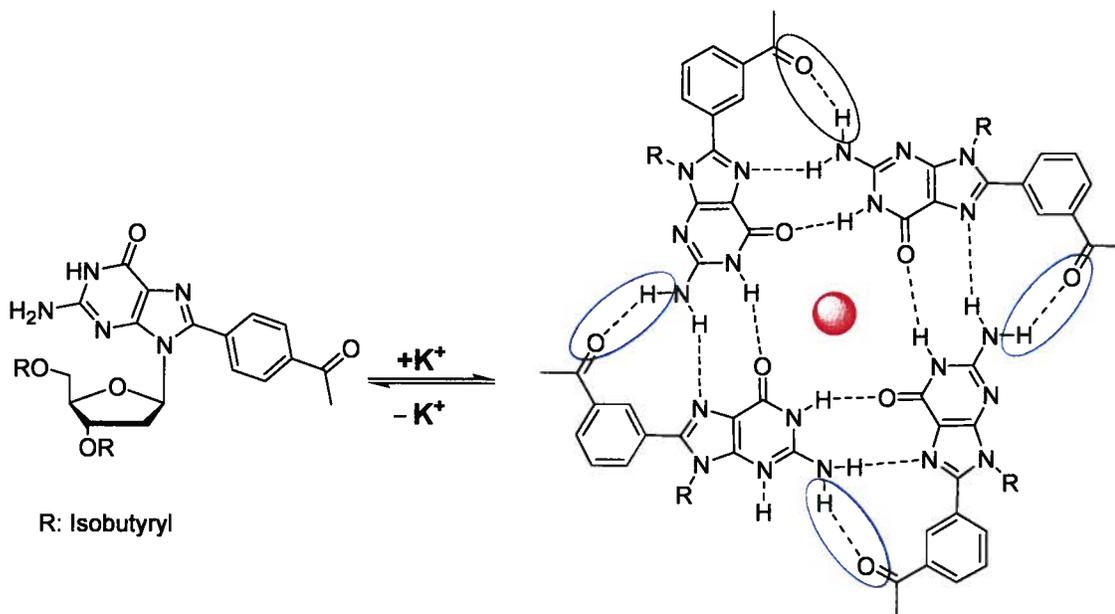


#### 1.4.7 A G-Quartet with Extra H-Bonds

In G-quartets, only one of the amino protons of the guanine participates in H-bonding. The other proton is either exposed to the solvent or bound to the anions.<sup>36</sup> Rivera and colleagues reported the synthesis of an 8-aryl-2'-deoxyguanosine compound (Figure 1.13) which was capable of forming extra H-bonds in the presence of  $K^+$ .<sup>51</sup> Variable-temperature  $^1H$  NMR and dilution experiments showed that the resulting G-quadruplex had considerable stability. Rivera et al. attributed the enhanced stability to a combination of factors including the formation of extra H-bonds and increased  $\pi$ -stacking interactions. The formation of a G-quartet from

modified guanine bases at N2 and C8 positions have also been reported by Wu's and Davis's groups.<sup>52,53</sup>

**Figure 1.13.** The G-quartet obtained from 8-aryl-2'-deoxyguanosine. The extra H-bonds are circled.



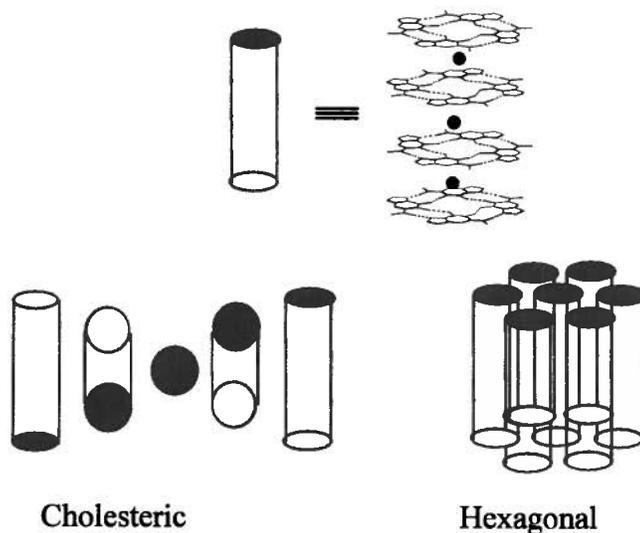
#### 1.4.8 G-Quartet in Material Science and Nanotechnology

G-quartets and G-quadruplexes owe their formation and properties to the unique structure of the guanine. The presence of the H-bond donor (NH and NH<sub>2</sub>) and acceptor (O6 and N7) faces along with its high self-association constants ( $K_{GG} \sim 10^2\text{-}10^4 \text{ M}^{-1}$  vs.  $K_{AA} < 5 \text{ M}^{-1}$  in CDCl<sub>3</sub>) and cation-coordination ability has made guanine an ideal candidate for the construction of supramolecular structures.<sup>54</sup> We give here a brief survey of the recent developments in this field. Although the main focus of this section will be on lipophilic systems, some examples from hydrophilic systems will also be presented in the appropriate place.

### 1.4.8.1 Liquid Crystalline Phases

As mentioned earlier, an octamer made from two stacked G-quartets is the basic observable unit of lipophilic G-quadruplexes in solution.<sup>29</sup> Under suitable conditions, not an octamer but a long columnar aggregate is the final product of the self-assembly.<sup>55</sup> Gottarelli and colleagues reported that the polymeric structures formed from lipophilic G-quartets exhibit lyotropic liquid crystalline properties (Figure 1.14).<sup>56</sup> Liquid crystal is a phase of matter which is somewhere between a liquid and a solid crystal.<sup>57</sup> The authors used small-angle neutron scattering (SANS) and NMR analysis to show that the aggregates are composed of G-quartets, regularly stacked on top of each other. Such behavior of lipophilic G-quadruplexes is very similar to the gel/liquid-crystal formation of 5'-GMP and some G-rich oligonucleotides in water.<sup>58,59</sup> Compounds such as folic acid which form tetrameric structures in water,<sup>60</sup> and G-ribbons<sup>18</sup> are also known to form liquid crystalline phases under suitable condition.

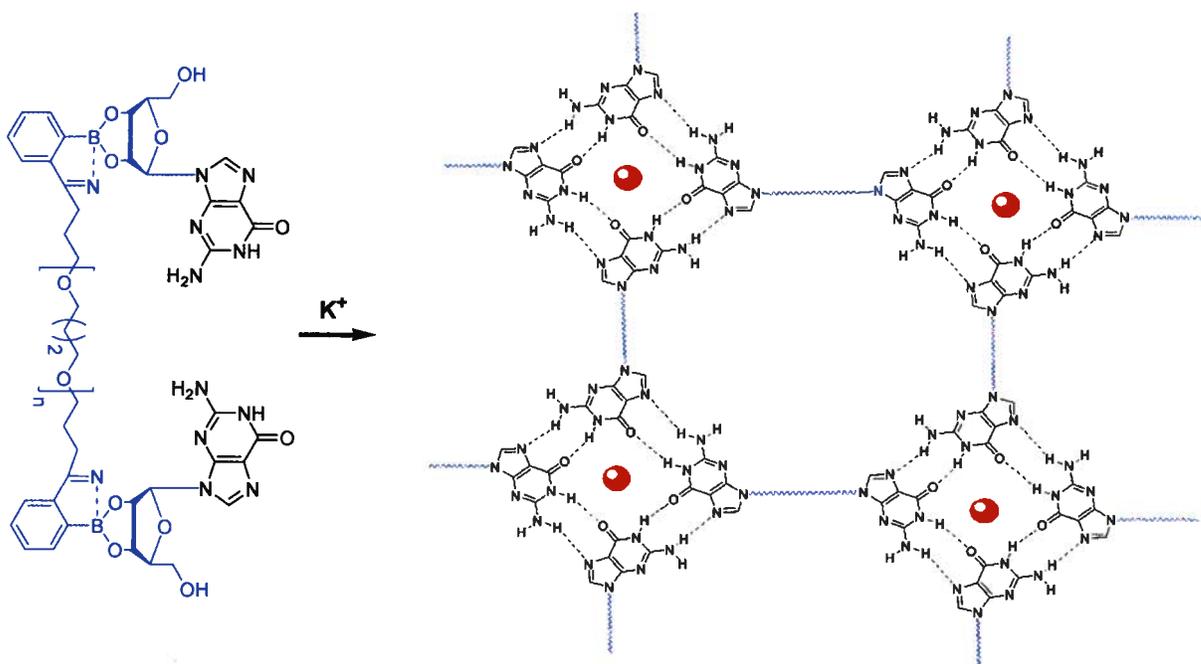
**Figure 1.14.** A schematic representation of the liquid crystalline phases obtained from lipophilic or hydrophilic guanosine compounds. Depending on concentration, ions, and temperature, either the cholesteric or the hexagonal mesophases is obtained.



### 1.4.8.2 Supramolecular Polymers

There have been a number of investigations aimed at the formation of polymeric films by making use of the structural capabilities of the nucleotides.<sup>61</sup> One commonly used strategy to achieve this is to functionalize the chain end of an appropriately chosen spacer with complementary bases. The predictable interaction of these recognition sites affords the supramolecular product.<sup>54,61</sup> This strategy has been recently employed by Barboiu et al. to prepare a G-quartet based membrane film.<sup>62</sup> Bisiminoboronate-guanosine (Figure 1.15), the symmetric monomer used in this study, forms a mixture of G-ribbons and other oligomers in the absence of cations, but solidifies upon addition of  $K^+$ . The resulting structure was found to be an ordered membrane film which allows rapid transport of electron/proton and  $Na^+/K^+$ .

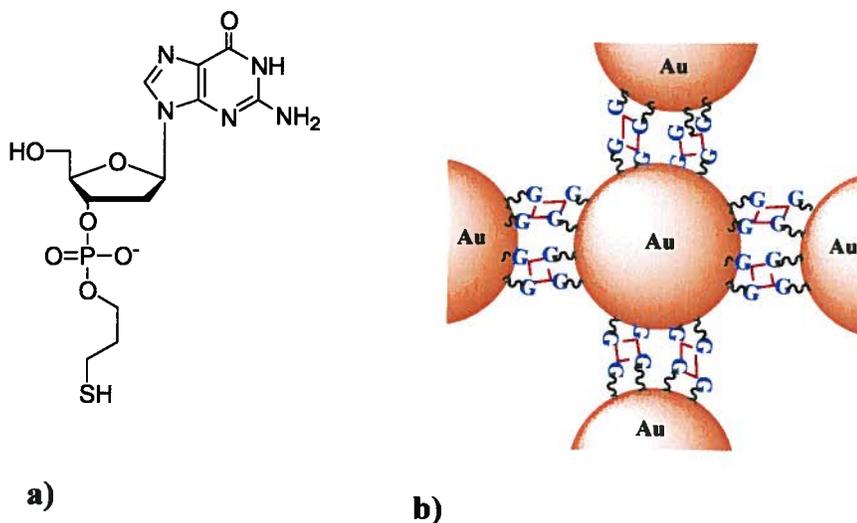
**Figure 1.15.** The self-assembly of bisiminoboronate-guanosine to a G-quartet-based polymeric film in the presence of  $K^+$ . The B-N dative bond of the monomer helps rigidify the polymer and also prevents the hydrolysis of the reversible boronate-guanosine ester bond.



### 1.4.8.3 Nanoparticle Assembly

Gaining control over the self-assembly of nanoparticles is essential for the formation of functional nanomaterials.<sup>63,64</sup> Mirkin and Li have reported the self-assembly of gold nanoparticles attached to the thiolated G-rich oligonucleotides.<sup>65</sup> In their experiment, 13 nm Au nanoparticles were first functionalized with thiolated guanosine derivative such as 3'-thiolpropyl deoxyguanosine phosphate (Figure 1.16) and were then exposed to cations. This led to the aggregation and consequent precipitation of the gold nanoparticles. Mirkin et al. have observed a cation-dependency for the self-assembly process ( $K^+ \gg Cs^+ > Na^+$ ) which suggests the involvement of G-quartet motifs. More recently, Shen and co-workers reported that the effect of cations on G-quartet induced aggregation of nanoparticle can be controlled by altering the length of the G-rich sequences.<sup>66</sup> It should also be noted that the non-specific interaction of DNA bases with gold nanoparticles is one of the challenges of this kind of studies.

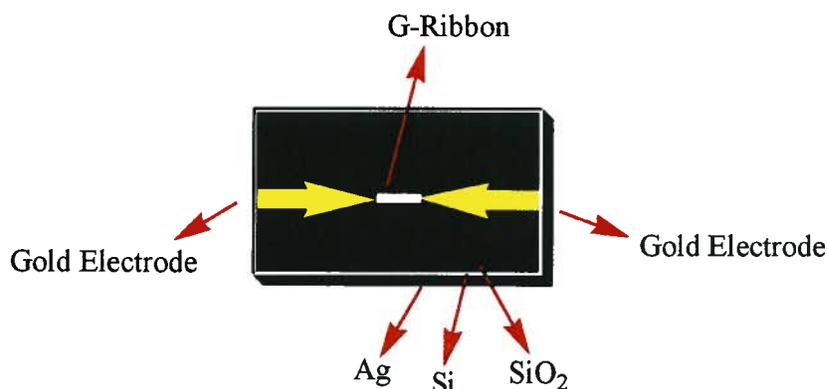
**Figure 1.16.** a) 3'-thiolpropyl deoxyguanosine phosphate and b) gold nanoparticle assembly using G-quartet formation (cations are omitted for clarity).



#### 1.4.8.4 Molecular Electronics

Biological molecules such as DNA have long been considered for the formation of bioelectronics.<sup>67,68</sup> Recently, it was established that double-stranded DNA is capable of charge transfer but is very sensitive to structural changes.<sup>69</sup> Guanine base itself has a low ionization potential and can assemble into an ordered structure in solution or on solid surfaces. Studies conducted by Calzolari et al.,<sup>70,71</sup> Barton<sup>72</sup> and Chatgililoglu et al.<sup>73</sup> have suggested that the structure of G-quartets could be used for the development of molecular electronics. However, there is still a need for more experimental work to support that. In comparison, more studies have been done on G-ribbons.<sup>74,75</sup> A study performed by Rinaldi and colleagues showed that the G-ribbons deposited between two gold electrodes on a mica surface (~100 nm away) are capable of conducting current (Figure 1.17, also see Figure 1.6c).<sup>74</sup> The current-potential curves (I-V) recorded by the authors showed a strong distance dependency. At short distances (~60nm) the device exhibited diode-like properties while at longer distances (~800 nm) it showed typical metal-semiconductor-metal properties. The random orientation of G-ribbons between the electrodes with respect to their overall dipoles was found to be a major problem for this study.

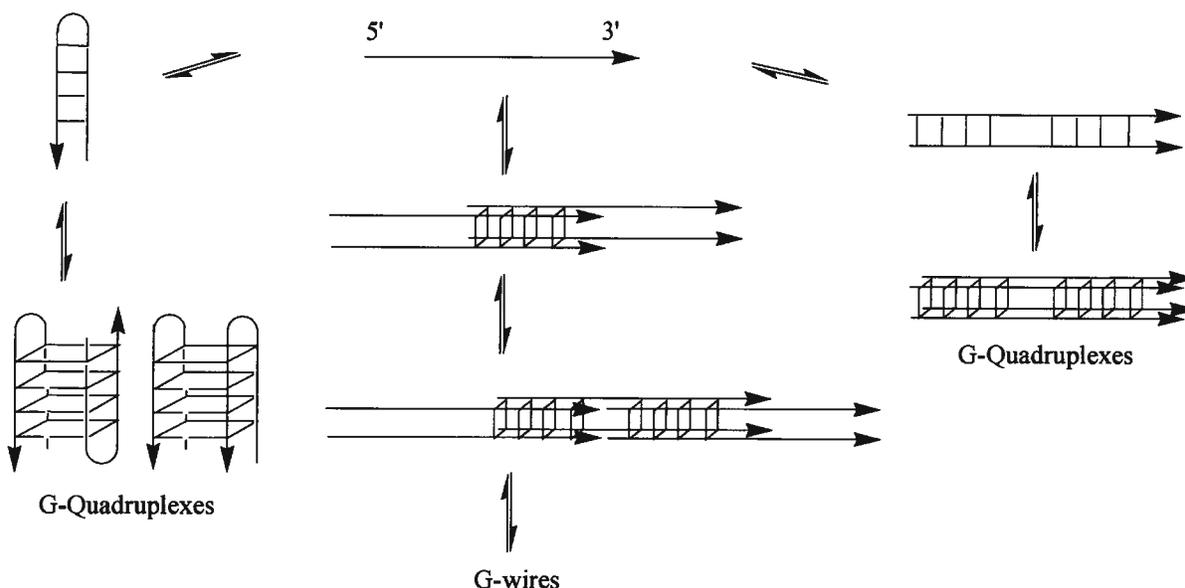
**Figure 1.17.** A schematic representation of the simple transistor made from G-ribbons.



### 1.4.8.5 Nanowires and G-Wires

DNA-based nanotubes and nanowires have been in the center of attention for their potential sensing or transport capabilities.<sup>76</sup> In the early 1990's Sen and Gilbert reported that DNA sequences with a 3'- or 5'-terminal guanine form high molecular weight structures in the presence of  $K^+$  which are distinct from smaller G-quadruplexes.<sup>77</sup> A few years later, Marsh and Henderson reported the AFM images of the polymers obtained from telomeric DNA  $d(G_4T_2G_4)$  in the presence of  $Mg^{2+}$  and termed them G-wires (Figure 1.18).<sup>78</sup> G-wires are linear and generally have very few bends, range from 10 nm to  $1\mu m$ , and have considerable thermodynamic stabilities. In contrast to double helices, the G-wires hold their shapes and structures when placed on a mica surface.<sup>79</sup> G-wires and their close relatives, the frayed wires<sup>80</sup>  $d(A_{15}G_{15})_n$  are currently under investigation as novel nanomaterials.

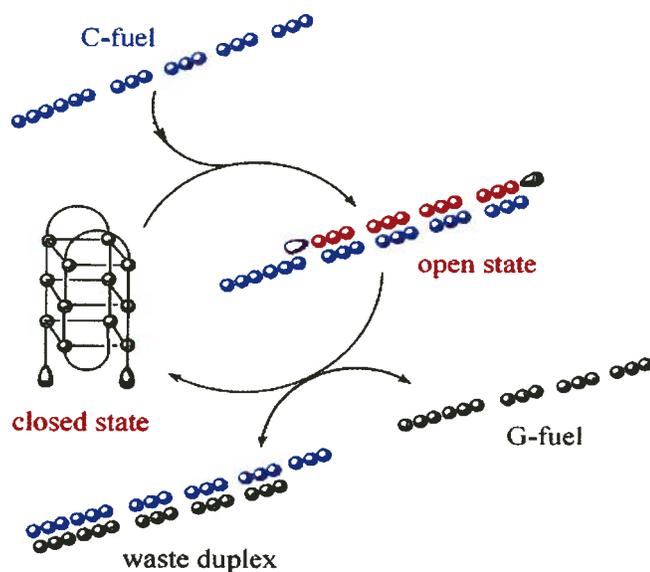
**Figure 1.18.** A schematic representation of the competing structures which can be adopted by telomeric DNA  $d(G_4T_2G_4)$ . The experimental conditions can be adjusted to favor the formation of G-wires. As shown below the formation of a G-wire is the result of the interaction between the 3' end of one helix and the 5' end of another helix (cations are omitted for clarity).



### 1.4.8.6 Nanomachines

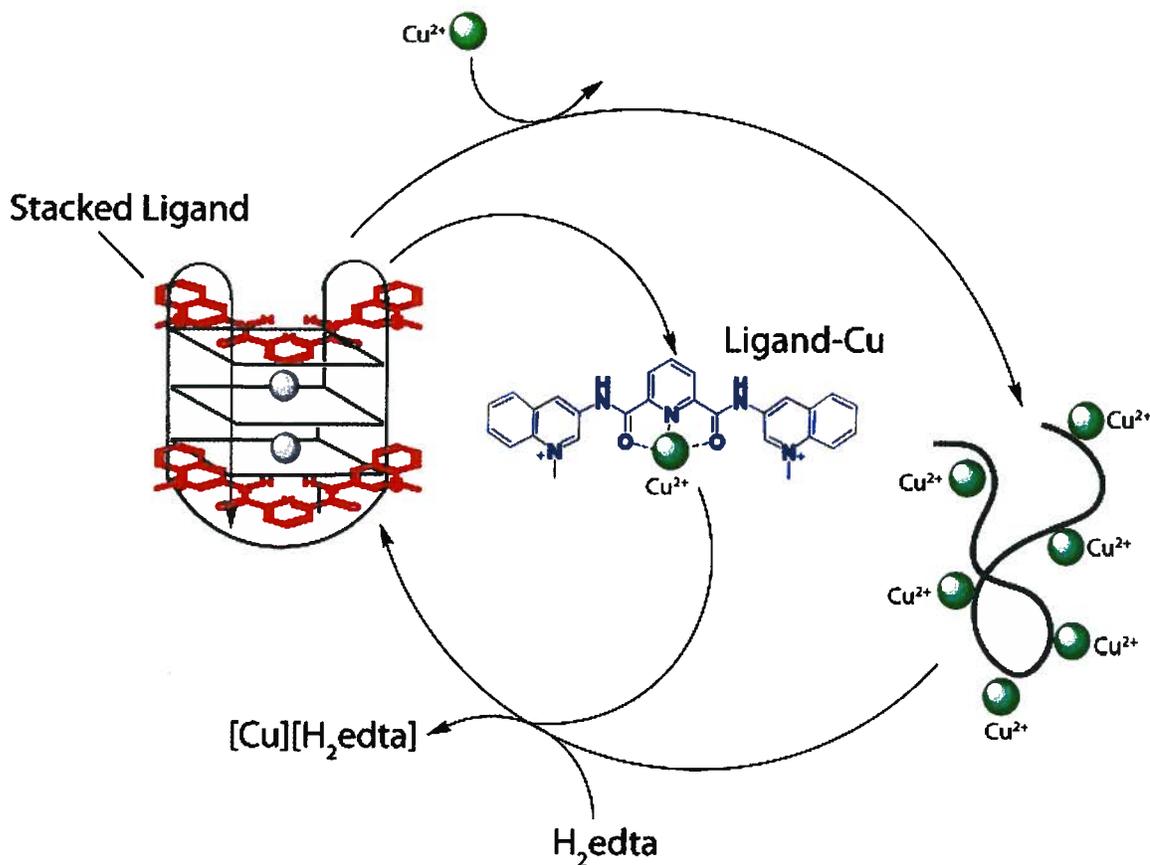
DNA-nanomachines have been inspired by biological molecular machines such as myosin and DNA-polymerase and are still in an early experimental stage. These systems are made by self-assembly and rely on sequence-specific interactions of DNA to perform a special task (e.g sensing, delivering or moving).<sup>81,82</sup> In 2002, Li and Tan reported that G-quadruplex-to-duplex transition, and vice versa, generates an extending-shrinking movement which can be the basis of a nanomotor.<sup>83</sup> Shortly thereafter, Alberti and Mergny employed a similar paradigm to construct a nanomachine based on the same G-quadruplex-to-duplex cycle.<sup>84</sup> As shown in Figure 1.19, the machine oscillates between an open and a closed state by the addition of two complementary strands (C-fuel and G-fuel) and generates a double-stranded waste. A 5'-fluorescein donor and a 3'-rhodamine acceptor were connected to the G-quadruplex to monitor the structural transition by FRET.

**Figure 1.19.** A schematic representation of a G-quadruplex-based nanomachine. The quadruplex-forming strand extends into a duplex upon addition of a complementary strand (C-fuel) and shrinks back to G-quadruplex via addition of G-fuel (cations are omitted for clarity).



In addition to duplex fueling, external parameters such as light or pH can also be used to trigger a conformational transition. Recently, Mergny et al. reported the copper-mediated transition of a G-quadruplex to a random coil using a bisquinolinium ligand.<sup>85</sup> The ligand can adopt two different conformations depending on whether it is free (V-shaped) or Cu-complexed (linear). Only the V-shaped conformation has affinity for the G-quadruplex. Mergny et al. designed a reversible cycle based on the copper-mediated transition of these two conformations. In the absence of copper the ligand is active and stabilizes the G-quadruplex structure. The addition of copper deactivates the ligand and at the same time denatures the DNA. Subsequent addition of EDTA traps the copper ion and regenerates the G-quadruplex. The whole process can be monitored by CD spectroscopy (Figure 1.20).

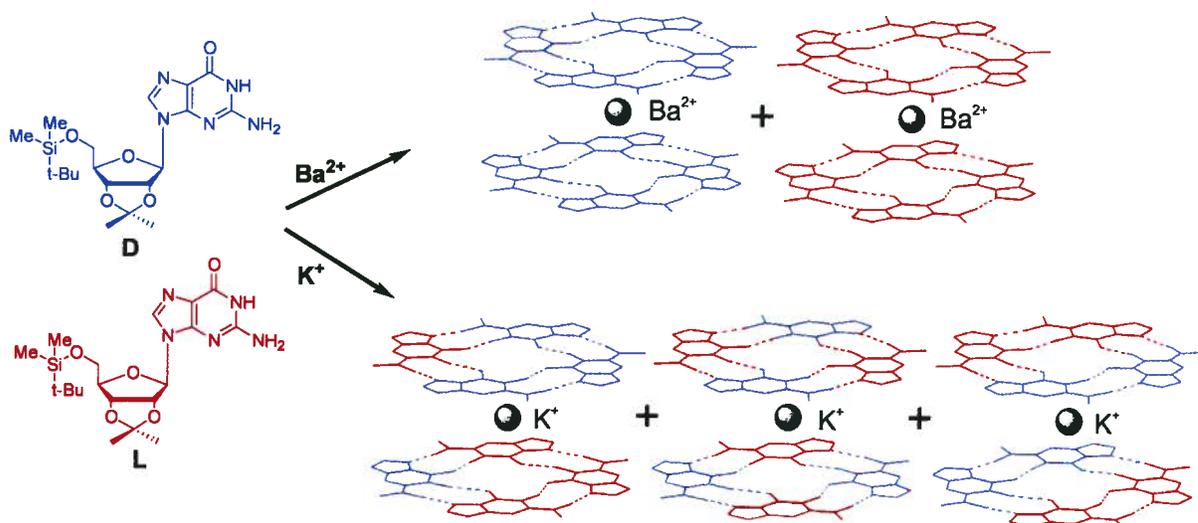
**Figure 1.20.** Copper-driven transition of a G-quadruplex to a random coil and vice versa.



### 1.4.9 G-Quartet and Chiral Resolution

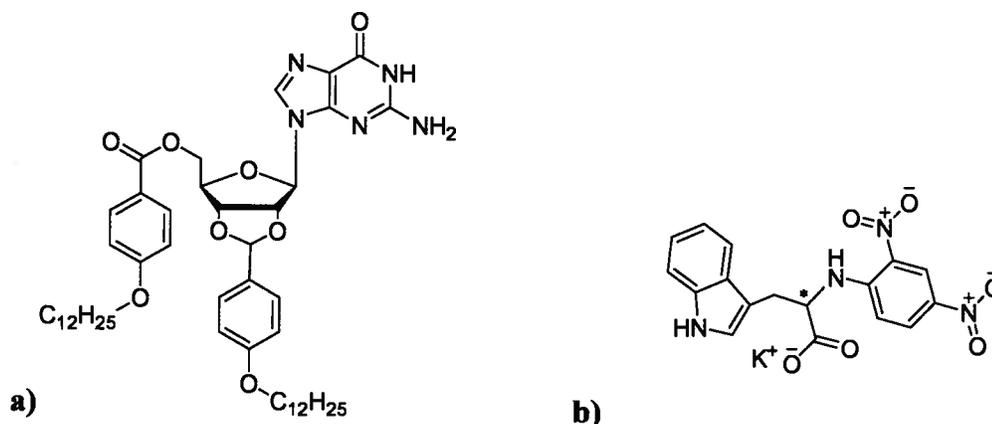
The importance of chirality at molecular level and in natural supramolecular systems (e.g. DNA and proteins) is clear.<sup>86</sup> Chiral supramolecular systems can be assembled from chiral or achiral building blocks. Guanosine compounds are chiral and the chirality information is contained in the ribose unit.<sup>87</sup> Davis's group showed that when a mixture of (*D*, *L*)-5'-silyl-2',3'-isopropylidene guanosine in CH<sub>2</sub>Cl<sub>2</sub> was mixed with an aqueous solution of barium picrate, only homochiral G-quadruplexes were formed (Figure 1.21).<sup>34</sup> In contrast, when potassium picrate was used for extraction, a mixture of heterochiral G-quadruplexes was obtained. Moreover, addition of Ba<sup>2+</sup> to the solution of K<sup>+</sup>-bound G-quadruplexes resulted in the formation of homochiral G-quadruplexes. Since Ba<sup>2+</sup> and K<sup>+</sup> have very similar sizes, they concluded that the divalent cation directs the enantiomeric self-recognition. They also suggested that the favorable enthalpy of Ba<sup>2+</sup>-binding is enough to overcome the unfavorable entropy of enantiomeric separation.

**Figure 1.21.** Homochiral and heterochiral G-quadruplexes obtained from extraction of (*D*, *L*)-5'-silyl-2',3'-isopropylidene guanosine with Ba<sup>2+</sup> and K<sup>+</sup> picrates.



In a related experiment, Gottarelli et al. showed that some lipophilic guanosine compounds can preferentially extract chiral amino acid anions from water into chloroform.<sup>88</sup> When the guanosine compound shown in Figure 1.22 was mixed with potassium salts of N-(2,4-dinitrophenyl)-(L,D)-tryptophan, a 3:1 enantioselectivity for (L) over (D)-enantiomer was observed. Although the enantiomeric excess observed for this process was not high (50 %), it was sufficient to prove the selective interaction of these systems with the chiral anions.

**Figure 1.22.** a) The lipophilic guanosine compound capable of chiral discrimination between (D) and (L) N-(2,4-dinitrophenyl)-tryptophan, and b) potassium salt of N-(2,4-dinitrophenyl)-(L,D)-tryptophan.



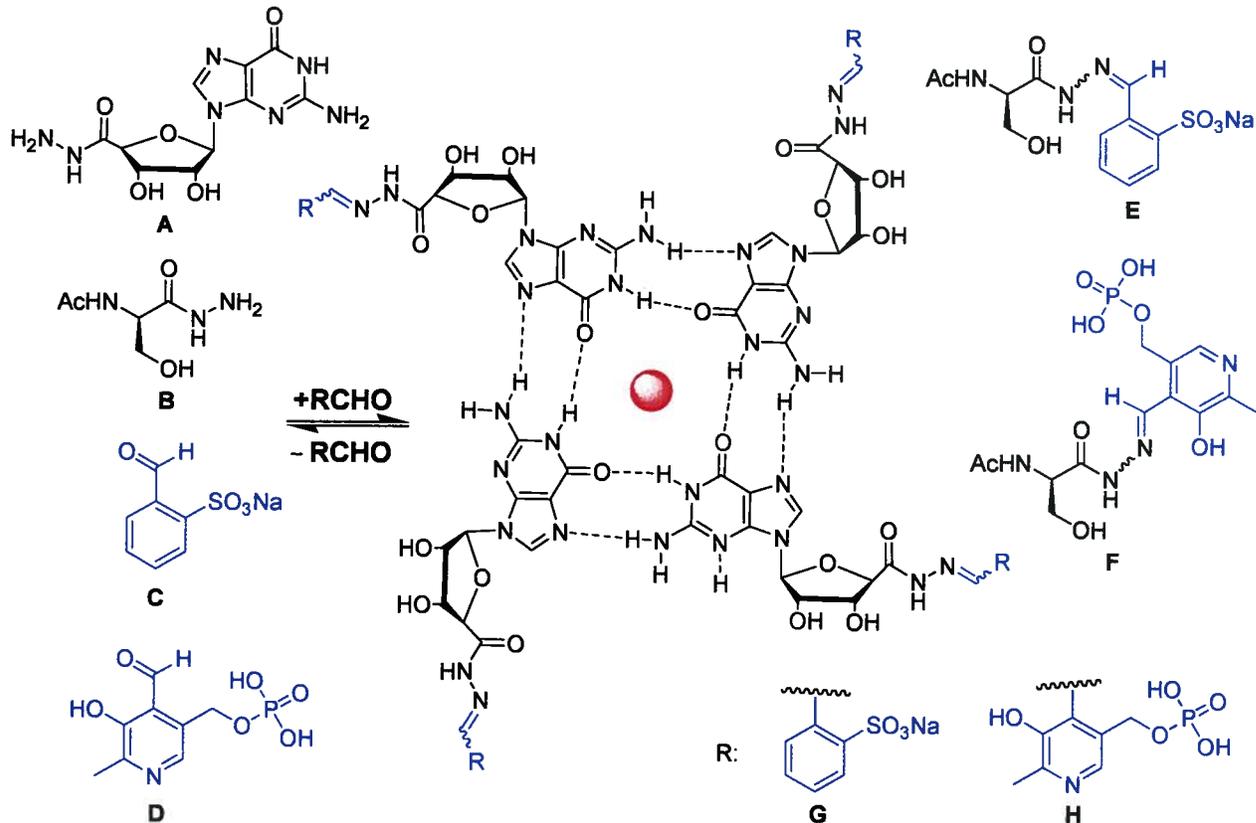
Besides recognition, guanosine compounds also offer potential for chiral separation. McGown et al. reported that the gels formed by 5'-GMP could be employed in capillary electrophoresis.<sup>89</sup> Using 5'-GMP as run buffer, an enantiomeric resolution of 2.1-2.3 for a mixture of (D, L)-propranol was obtained. The authors also suggested that the reversible nature of gelation could facilitate the final separation of the purified enantiomers from the gel.

## 1.4.10 G-Quartet and Equilibriums

### 1.4.10.1 Dynamic Covalent Chemistry (DCC)

Dynamic covalent chemistry enables the generation of a library of compounds at equilibrium from which molecules with interesting properties can be selected by either physical (e.g. temperature) or chemical (e.g. pH) means.<sup>90,91</sup> Lehn and Sreenivasachary reported the formation and isolation of a stable acylhydrazone gel from the reaction of guanosine hydrazone and various aldehydes.<sup>92</sup> Hydrazides (A and B) and aldehydes (C and D) were reacted and the distribution of the products was then investigated by <sup>1</sup>H NMR spectroscopy (Figure 1.23).

**Figure 1.23.** Preparation of a library of acylhydrazones (E, F) and G-quartet hydrazones (G, H) from the reaction of guanosine hydrazone A and serine hydrazone B with aldehydes C and D.

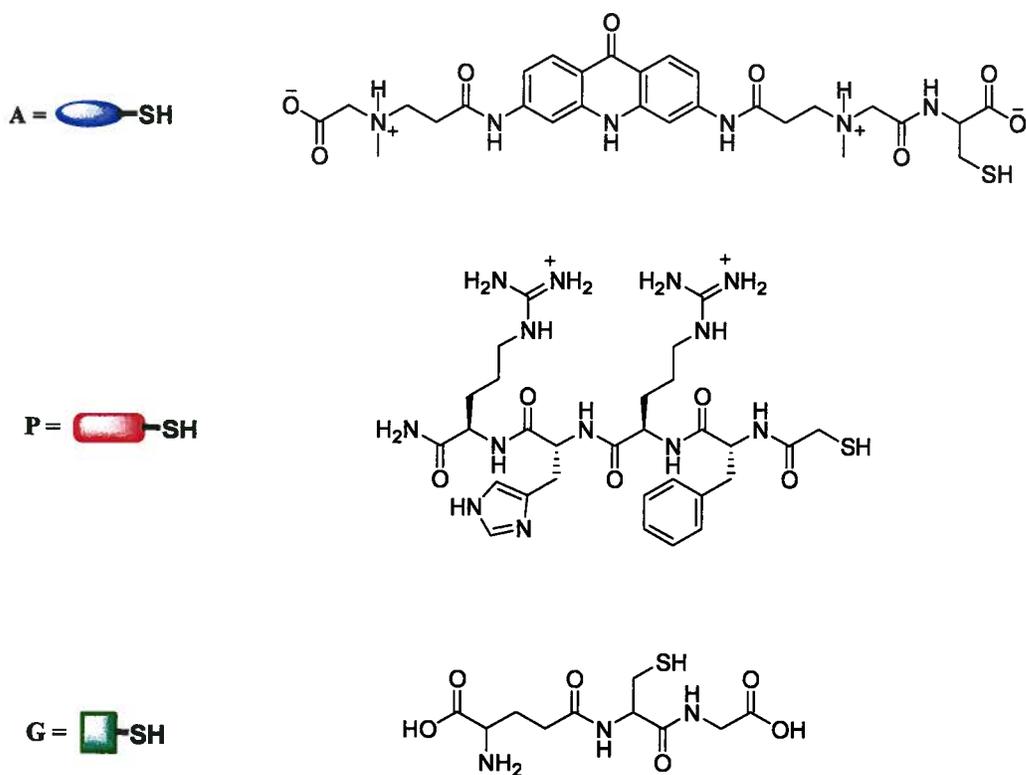
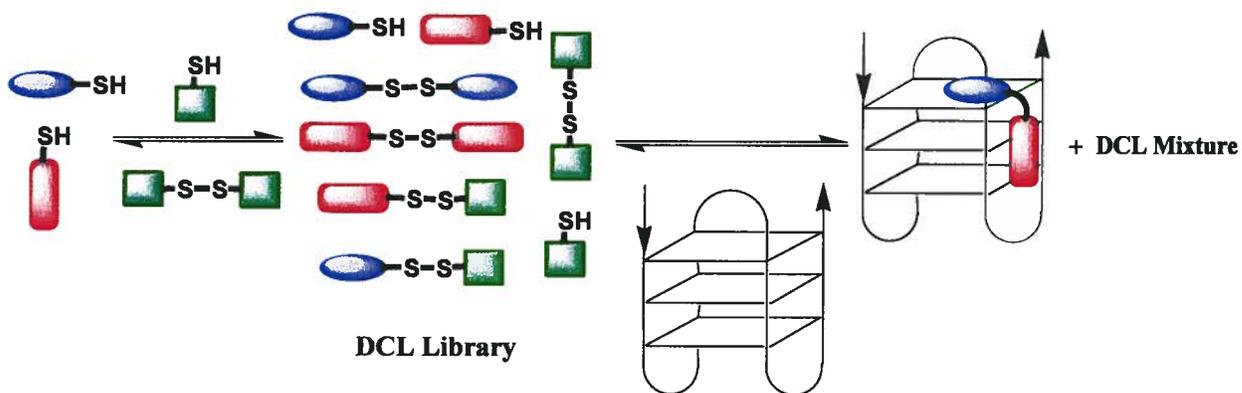


The gelation process was found to be thermoreversible. At 80 °C, a temperature which was not favorable for G-quartet formation, only a uniform mixture of acylhydrazones was obtained containing all possible products (E, F, G, and H) at equilibrium. At 25 °C, the distribution of products was uneven and two acylhydrazones H and E predominated in the gel and solution respectively. These findings clearly indicate that the gelation process can be used as a powerful selection tool in thermodynamically controlled reversible reactions. More recently, Lehn and Sreenivasachary have shown that the guanosine hydrazides gels can be promising media for controlled drug delivery.<sup>93</sup>

The impact of G-quartet formation on equilibrating systems has also been exploited in the field of nucleic acid G-quadruplexes. Balasubramanian and colleagues have performed a number of studies using G-quadruplexes as templates for ligand discovery.<sup>94-96</sup> As mentioned earlier, ligands, which can selectively stabilize or interact with nucleic acid G-quadruplexes, are of interest and under extensive investigation.<sup>24,25</sup> Such ligands are generally composed of two domains: 1) hydrophobic cores which are capable of  $\pi$ -stacking and intercalation, and 2) charged arms which are capable of interacting with grooves or loops. In one study, Balasubramanian et al. generated a library of two ligands which were capable of interacting with G-quadruplexes only via one of these domains. These ligands, a  $\pi$ -stacking acridone derivative (A) and a groove-binding peptide (P), were functionalized with thiol groups and kept at pH = 7.4 in the presence of glutathione (G) (Figure 1.24). This condition allowed formation and rapid exchange of disulfide bonds among the reactants. In the absence of G-quadruplex the reaction was under thermodynamic control and a mixture of homo or heterodisulfides were obtained. Upon addition of G-quadruplex, a new equilibrium was established and a four times increase in the molar ratio of A-P, a ligand containing both of the domains, was observed (Figure 1.24). Interestingly, the formation of peptide disulfide (P-P) was also amplified (five times) using the G-quadruplex

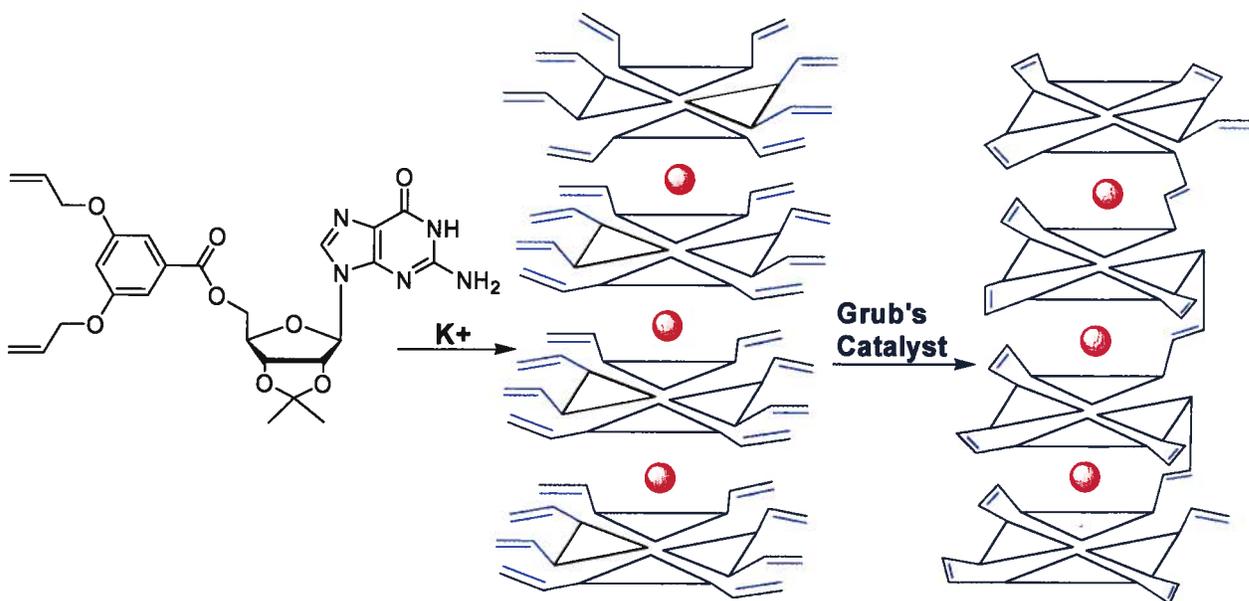
template. Further experiments showed that the newly synthesized ligands had higher association constant with the G-quadruplex.<sup>94</sup>

**Figure 1.24.** A schematic representation of the reaction of acridone (A), peptide (P), and glutathione (G, both oxidized and reduced forms) in the presence and absence of the G-quadruplex template.



Davis and co-workers reported the formation of an artificial ion channel on the basis of DCC.<sup>97</sup> Ion channels are pore-forming proteins that regulate the transport of ions, particularly sodium and potassium, across the cell membrane. In recent years, there have been a number of investigations to build artificial ion channels. Such models are generally formed from synthetic ionophores and need to be selective for a particular cation.<sup>98</sup> Lipophilic G-quadruplexes are often cited as one of these models owing to their shape and selectivity for cations.<sup>30</sup> In spite of their thermodynamic stabilities, lipophilic G-quadruplexes lack the kinetic stability to serve as ion channels.<sup>97</sup> Using reversible olefin metathesis, Davis et al. succeeded in cross-linking the guanosine units within a G-quadruplex to form a kinetically stable unimolecular G-quadruplex (Figure 1.25). The corresponding G-quadruplex was found to be capable of transporting sodium ions across a liposomal membrane. The direct evidence of sodium transport was obtained from <sup>23</sup>Na analysis.

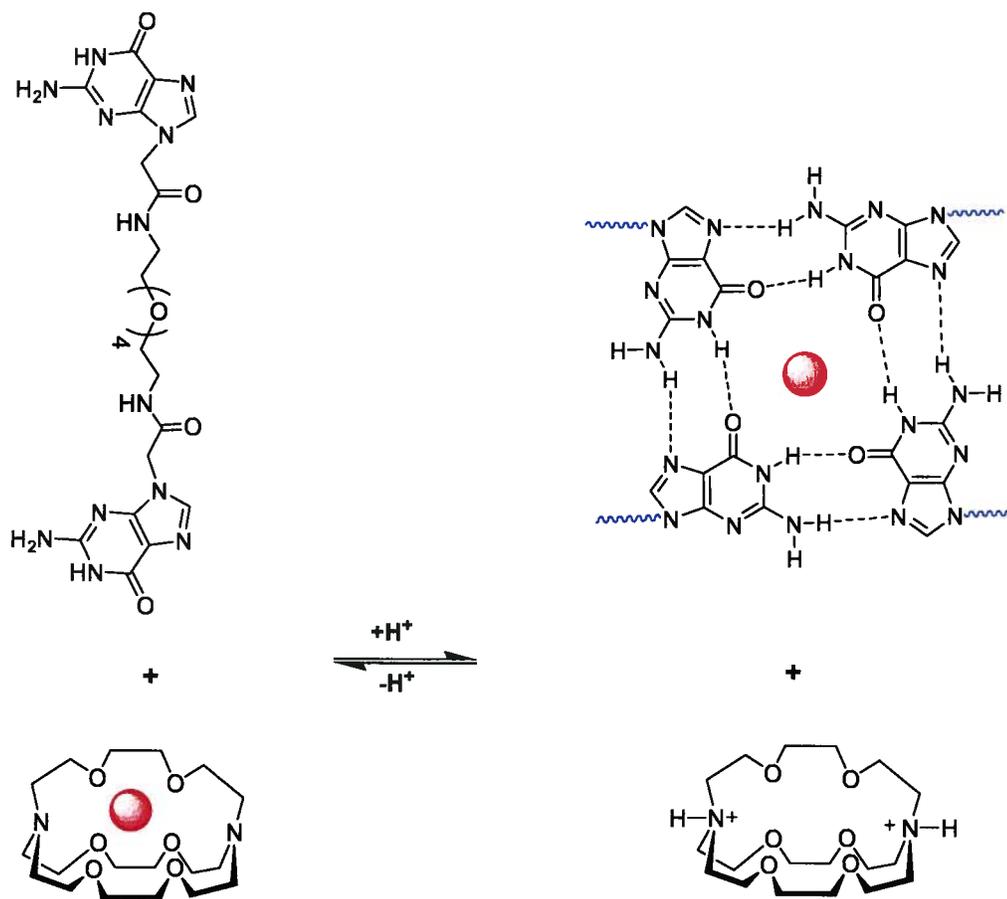
**Figure 1.25.** Synthesis of a unimolecular G-quadruplex using reversible olefin metathesis.



### 1.4.10.2 Dynamic Cation Binding and Release

As shown in section 1.4.8.2, homoditopic guanine-terminated monomers can form stable polymeric networks in the presence of cations. Ghossoub and Lehn reported the sol-gel transition of a bisguanine compound via the binding and release of  $K^+$  by [2.2.2] cryptand in water (Figure 1.26).<sup>99</sup> Cryptand binds  $K^+$  better than guanine and causes a shift in the equilibrium towards the starting materials. The gel-state can be regenerated again by protonation of the bridgehead nitrogens. In a similar study, Spada et al. reported the interconversion of an octamer to a G-ribbon in chloroform using [2.2.2] cryptand.<sup>100</sup>

**Figure 1.26.** Interconversion of a bisguanine compound between sol and gel states.



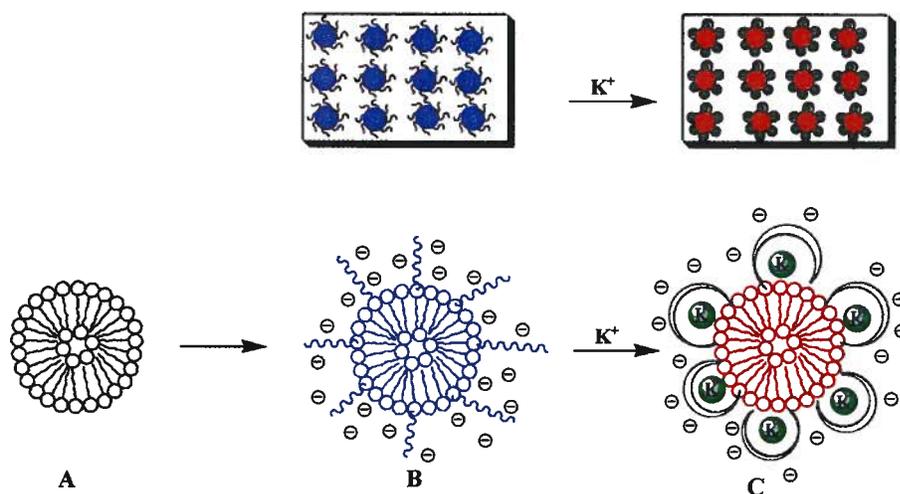
### **1.4.11 Nucleic Acid G-Quadruplexes as Probes**

Nucleic acid G-quadruplexes are significant both as a target and as a probe. Up to this point, we have reviewed some examples where G-quadruplexes were the target of small ligands mainly for therapeutic purposes. G-quadruplexes also have the potential to serve as biosensors for bioanalytical applications due to their thermodynamic stability and specific interactions.

#### **1.4.11.1 G-Quadruplexes as Cation-Sensors**

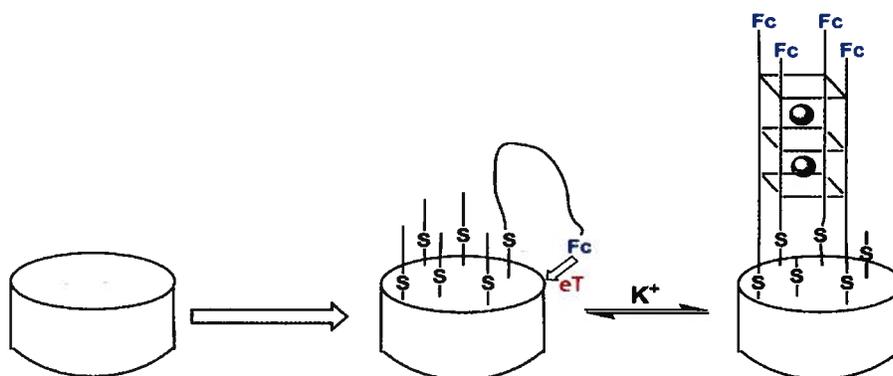
Cation-binding has an important impact on the structure of G-quadruplexes. Thus, a number of biosensors have been developed that utilize the cation-G-quadruplex interactions as a driving force. Kim and co-workers designed a calorimetric potassium biosensor on the basis of the higher selectivity of a G-quadruplex for potassium over sodium.<sup>101</sup> The physiological concentration of potassium in blood (3.5-5.3 mM) is significantly lower than that of sodium (135-145 mM). Thus, selective detection of potassium in the presence of sodium is a challenging task. The system developed by Kim et al. comprised a microarray of polydiacetylene (PDA) liposomes functionalized with G-rich oligonucleotides. The G-rich oligonucleotides were distributed on the surface of the liposome in a way to form a densely packed layer. PDA is not fluorescent under normal conditions but undergoes a color change when is subjected to mechanical stress. Upon recognition of potassium, bulky quadruplexes are formed that repel each other and impose a mechanical stress on the PDA backbone. This activates the PDA and a red fluorescence is observed (Figure 1.27). CD spectroscopic experiments confirmed the formation of the G-quadruplexes.

**Figure 1.27.** A) polydiacetylene (PDA) liposome, B) a polydiacetylene (PDA) liposome functionalized with G-rich oligonucleotides (non fluorescent) and the corresponding microarray, and C) activated liposome (red fluorescent) due to the formation of G-quadruplex.



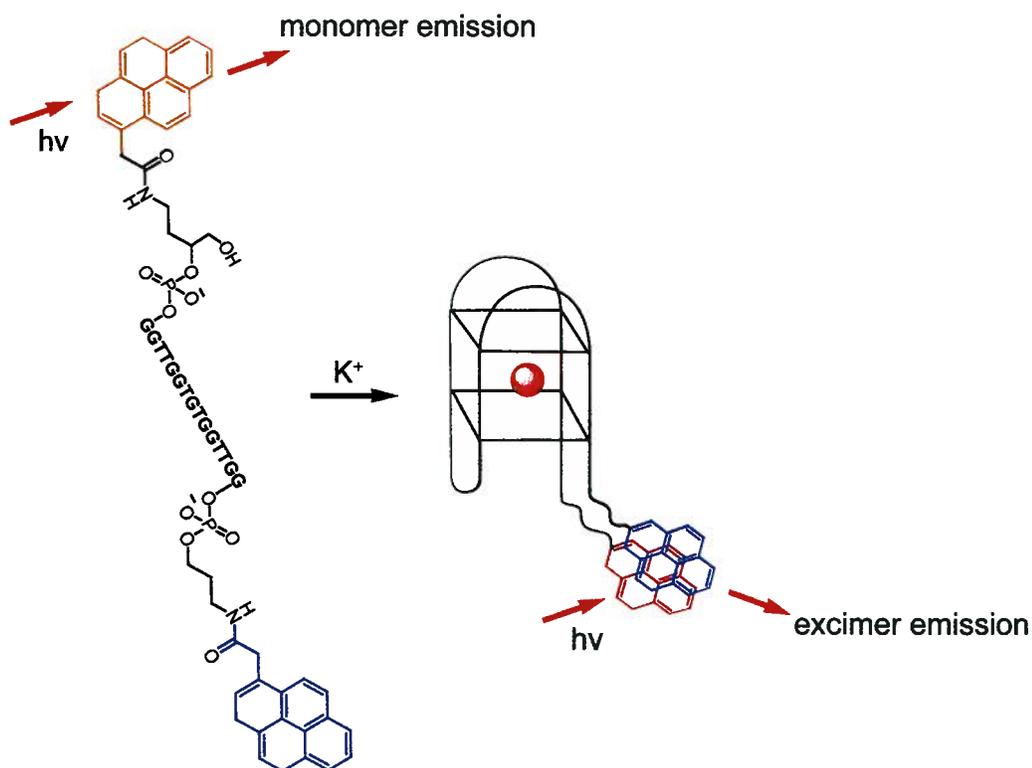
More recently, Wu et al. developed an electrochemical nanoswitch for the detection of  $K^+$  using G-rich strands immobilized on a gold electrode (Figure 1.28).<sup>102</sup> The oligonucleotides were labeled with a redox-active ferrocene derivative (Fc) which could only interact with the electrode in the absence of  $K^+$  (switch-on). In the presence of  $K^+$ , a tetramolecular G-quadruplex was formed and the electron transfer was terminated (switch-off). The redox reaction and the on/off states were monitored by AC voltammetry and the nanoswitch could be regenerated.

**Figure 1.28.** A schematic representation of the on/off nanoswitch used for the electrochemical detection of potassium ion.



G-quadruplex-based nanoswitches have also been reported by Radi, O'Sullivan<sup>103</sup> and Plaxco et al.<sup>104</sup> An optical-sensor was developed by Takenaka and co-workers for potassium-sensing (Figure 1.29).<sup>105</sup> The main skeleton of the sensor was a G-rich sequence d(5'-GGTTGGTGTGGTTGG-3') known as thrombin-binding aptamer (TBA) that arranges into a chair-type conformation in the presence of thrombin (blood-clotting protein).<sup>106</sup> Two pyrene moieties were incorporated as fluorophores to the termini of the TBA. In the absence of  $K^+$ , the monomer emission ( $\sim 390$  nm) was dominant while potassium-binding brought the pyrene moieties together and induced an excimer emission ( $>450$  nm). The system has shown a good selectivity for potassium over sodium at extracellular concentrations. Similar systems have been developed for the detection of other cations such as  $Ca^{2+}$  and  $Mg^{2+}$ .<sup>107</sup>

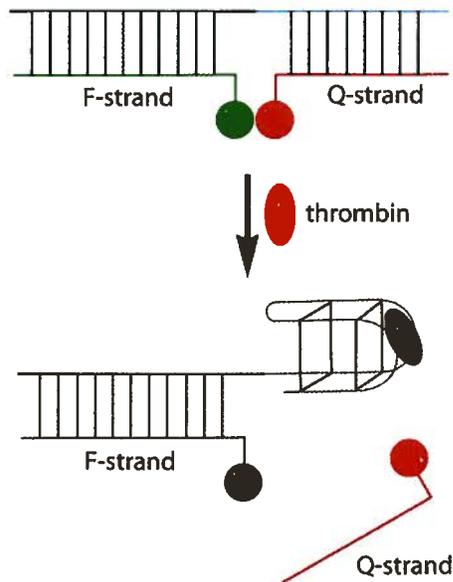
**Figure 1.29.** The structure of the thrombin-binding aptamer containing two pyrene moieties and the resulting chair-type conformation. The interaction of the fluorophores is only possible in the presence of potassium ions.



### 1.4.11.2 G-Quadruplexes as Protein-Sensors

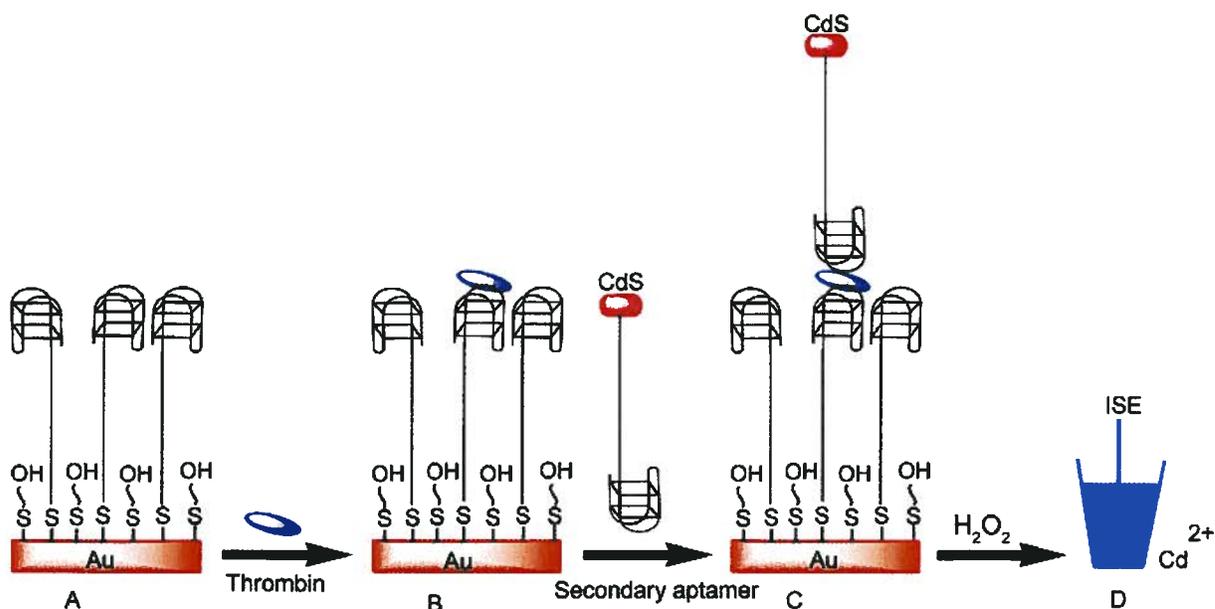
Rapid, low-cost detection of proteins is important for therapeutic purposes or for monitoring the cellular events. Early studies on quadruplex-based sensors were conducted mostly on the thrombin-binding aptamer.<sup>108</sup> Nutiu and Li designed a three-component FRET-based system for the detection of thrombin.<sup>109</sup> The strands were complementary to each other and designed in a way to form a duplex. The first strand (F-strand) was attached to a fluorophore at the 5'-end (green); the second strand (Q-strand) was modified with a quencher at the 3'-end (red) and the third strand contained the thrombin-binding sequence (blue) and was complementary to Q and F. Upon addition of thrombin, the thrombin-binding aptamer (blue) interacts with this protein and forms a G-quadruplex. Consequently, the Q-strand (red) is released and a strong fluorescence is observed (Figure 1.30).

**Figure 1.30.** Schematic illustration of the three-component duplex. Addition of thrombin leads to the formation of the G-quadruplex, which coincides with the release of the Q-strand and an enhancement in the fluorescence.



Aptamer-based sensors for the detection of thrombin have also been reported by other groups such as Heyduk et al.,<sup>110</sup> Chang et al.,<sup>111</sup> Yan et al.,<sup>112</sup> and Willner et al.<sup>113</sup> In recent years, a great deal of attention has been driven towards the development of electrochemical sensors due to their lower cost, fast response, and easy handling.<sup>114</sup> Wang, Bakker and Pretsch have introduced a potentiometric method for the detection of thrombin using ion-selective electrodes (ISEs).<sup>115</sup> First, the thrombin-binding aptamers were adsorbed on the surface of a gold electrode through the Au–S linkages. Then, the target protein, bound to a G-quadruplex, was added. The gold electrode was then treated with secondary aptamers labeled with a CdS nanoparticle. Finally, the CdS was oxidized (solubilized) with H<sub>2</sub>O<sub>2</sub> and detected potentiometrically with a Cd<sup>2+</sup>-selective microelectrode. A low detection limit of 5 ppb has been observed for this method making it an ideal method for the detection of proteins in low concentrations and volumes (μL). Aside from protein-detection, G-quadruplex-forming aptamers have also been utilized in electrophoresis for the isolation of peptides.<sup>116</sup>

**Figure 1.31.** A) Formation of monolayer, B) addition and capture of thrombin, C) addition and binding of the secondary aptamer, and D) oxidation and detection of Cd<sup>2+</sup> with an ion-selective electrode.

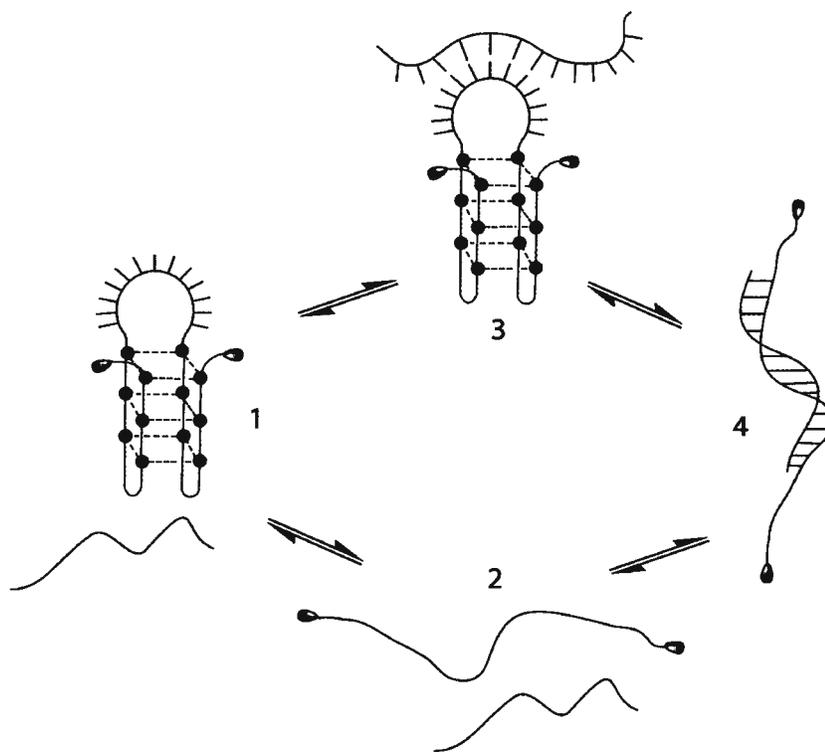


### 1.4.11.3 G-Quadruplexes as Oligonucleotide-Sensors

There are various methods for the detection of nucleic acids including UV spectroscopy and staining with ethidium bromide in conjunction with electrophoresis. However, only a few of these methods (e.g. southern blot) are sequence-specific. Analytical biosensors, such as the one that will be described in this section, allow for a sequence-dependent response which is needed for analytical works.

Gosse, Jullien, and Mergny designed a G-quadruplex-based molecular beacon to detect a target oligonucleotide with high efficiency.<sup>117</sup> Molecular beacons (MBs) are hairpin-shaped oligonucleotide probes that become fluorescent upon binding to complementary sequences.<sup>118</sup> Moreover, they are highly specific probes that can discriminate between target molecules with even one single mismatch.<sup>118</sup> While beacons are typically formed by Watson-Crick pairing, Gosse et al. showed that they can also be formed from G-quadruplexes with a central loop (dodecamer).<sup>117</sup> As shown in Figure 1.32, the addition of a complementary sequence results in the opening of the G-quadruplex and the formation of a duplex which in turn is accompanied by a change in the fluorescence emission. This step is similar to the duplex-quadruplex equilibria previously described for the nanomachines (section 1.4.8.6). According to kinetic studies, two competing mechanisms were proposed for this structural transition. In the first pathway, partial recognition of the beacon by the target leads to the opening of the G-quadruplex; in the second pathway, however, spontaneous opening of the beacon precedes the recognition stage (Figure 1.32). The presence of a single base mismatch has also been shown to prevent the opening of the G-quadruplex, demonstrating the high sequence-specificity of the probe.

**Figure 1.32.** Proposed mechanisms for the recognition of the G-quadruplex-based molecular beacon.

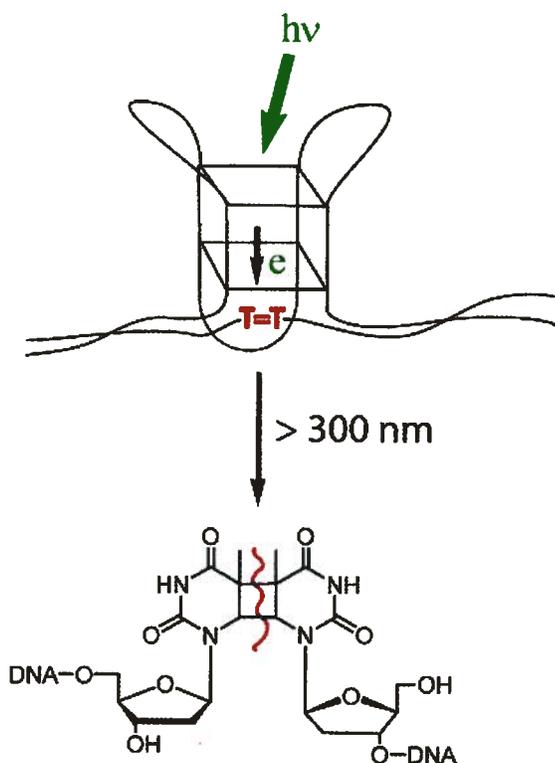


#### 1.4.12 G-Quadruplexes as Catalysts

Unlike RNA molecules whose catalytic activity is well established (ribozymes), DNA is only associated with the replication of genetic information.<sup>119</sup> It was shown in 1994 that single-stranded DNA is capable of accelerating some reactions *in vitro*.<sup>120</sup> Since then, numerous reports have appeared in the literature describing a variety of DNA-catalyzed reactions.<sup>119</sup> Catalytic activity of G-quadruplex has also been evaluated by several research groups.<sup>121-123</sup> Sen and Cinnapen reported that a deoxyribozyme containing a short quadruplex could cleave a thymine cyclobutane dimer when irradiated by near-UV light at 305 nm (Figure 1.33).<sup>124,125</sup>

They showed that the G-quadruplex could act as a catalytic domain that absorbs and transmits the low-energy radiation to the bound substrate. Interestingly, the absorbance and transmittance of energy was originally designed to be carried out in the presence of serotonin as a cofactor, but further experiments showed that the G-quadruplex was indeed the light-harvesting antenna.

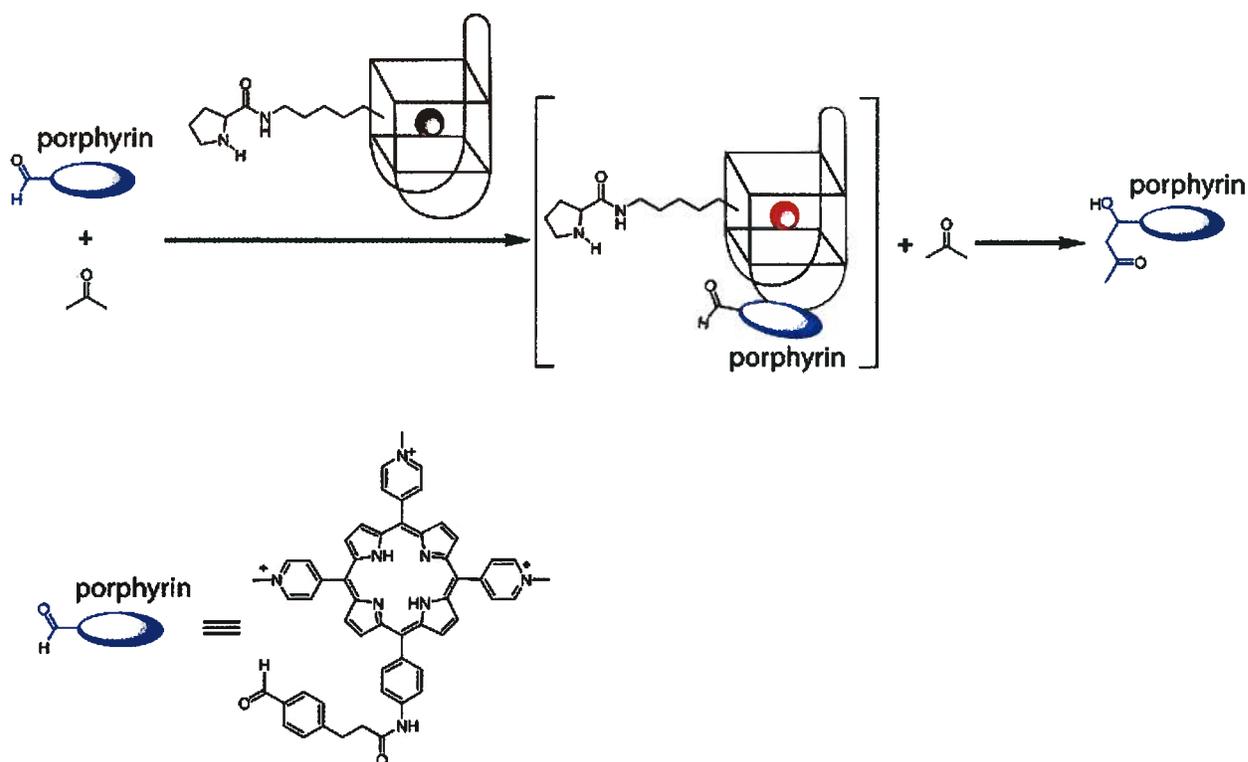
**Figure 1.33.** A schematic representation of the deoxyribozyme catalyzing the photocleavage of a thymine cyclobutane dimer. The G-quadruplex absorbs and transmits the low-energy radiation ( $\sim 305$  nm) which is outside the normal absorption of DNA ( $>250$  nm).



In a different set of studies, Sen and Li reported that guanine-rich oligomers could catalyze the metallation of porphyrins by copper and zinc ions.<sup>123</sup> More recently, Hartig, Marx and colleagues showed that proline-modified G quadruplexes could catalyze the aldol reaction between acetone and an aldehyde-appended porphyrin (Figure 1.34).<sup>126</sup> Cationic porphyrins such as tetramethylpyridinium porphyrin (TMPyP4) have been shown to bind to G-quadruplexes

and stabilize their structure.<sup>127</sup> Thus, the interaction of porphyrins with G-quadruplexes positions the aldehydes in close proximity to the catalytic site and enhances the rate of the reaction. The G-quadruplex seems to only template the reaction and does not participate directly in the catalysis. Using this strategy, the authors have observed a rate enhancement of more than three fold.

**Figure 1.34.** A proline modified G-quadruplex catalyzes the aldol reaction between acetone and an aldehyde-appended porphyrin.

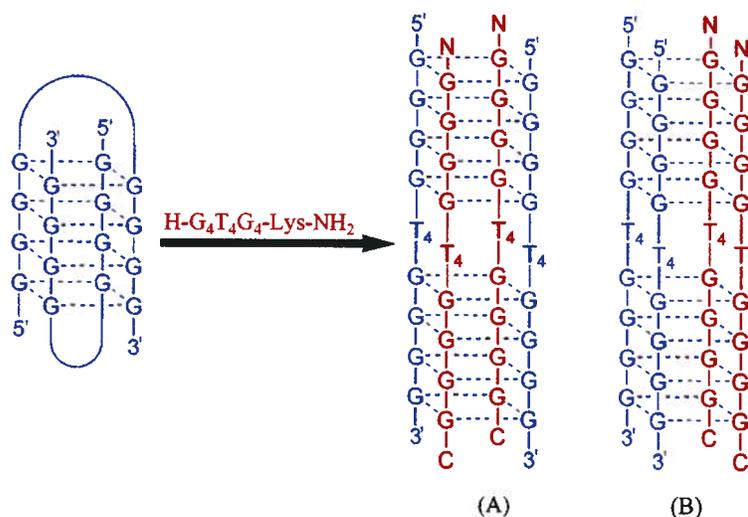


### 1.4.13 Miscellaneous G-Quadruplex-Based Structures

#### 1.4.13.1 Peptide Nucleic Acid (PNA) G-Quadruplex

Peptide-nucleic acids are synthetic polymers in which the sugar-phosphate backbone has been replaced by a peptide unit.<sup>128</sup> PNA are of interest in biological research due to their ability to bind to complementary DNA.<sup>128</sup> Armitage et al. showed that the base-pair recognition of DNA/PNA can also be applied to G-quadruplex structures.<sup>129</sup> They showed that a bimolecular G-quadruplex formed by the telomeric sequence of d(G<sub>4</sub>T<sub>4</sub>G<sub>4</sub>) is disrupted in the presence of PNA (H-G<sub>4</sub>T<sub>4</sub>G<sub>4</sub>-Lys-NH<sub>2</sub>) to form a hybrid quadruplex (PNA<sub>2</sub>-DNA<sub>2</sub>) (Figure 1.35).

**Figure 1.35.** Hybrid tetramolecular G-quadruplexes formed from DNA (blue) and PNA (Red). Two structures of diagonally opposite (A) or adjacent (B) are possible.



FRET experiments revealed that the PNA strands are parallel and that the 5'-termini of DNA align with the N-termini of the PNA. Two possible structures, (A) and (B), were proposed for the products. Of these, structure (A) was shown to be the real structure based on the FRET

results. The latter structure was also speculated to provide a better charge separation. In a similar study, Balasubramanian and co-workers reported the formation of a G-quadruplex composed entirely of PNA (pGT<sub>3</sub>).<sup>130</sup> UV-melting experiments showed that the resulting PNA tetraplex is less stable than its analogous DNA G-quadruplex d(G<sub>3</sub>T)<sub>4</sub>.

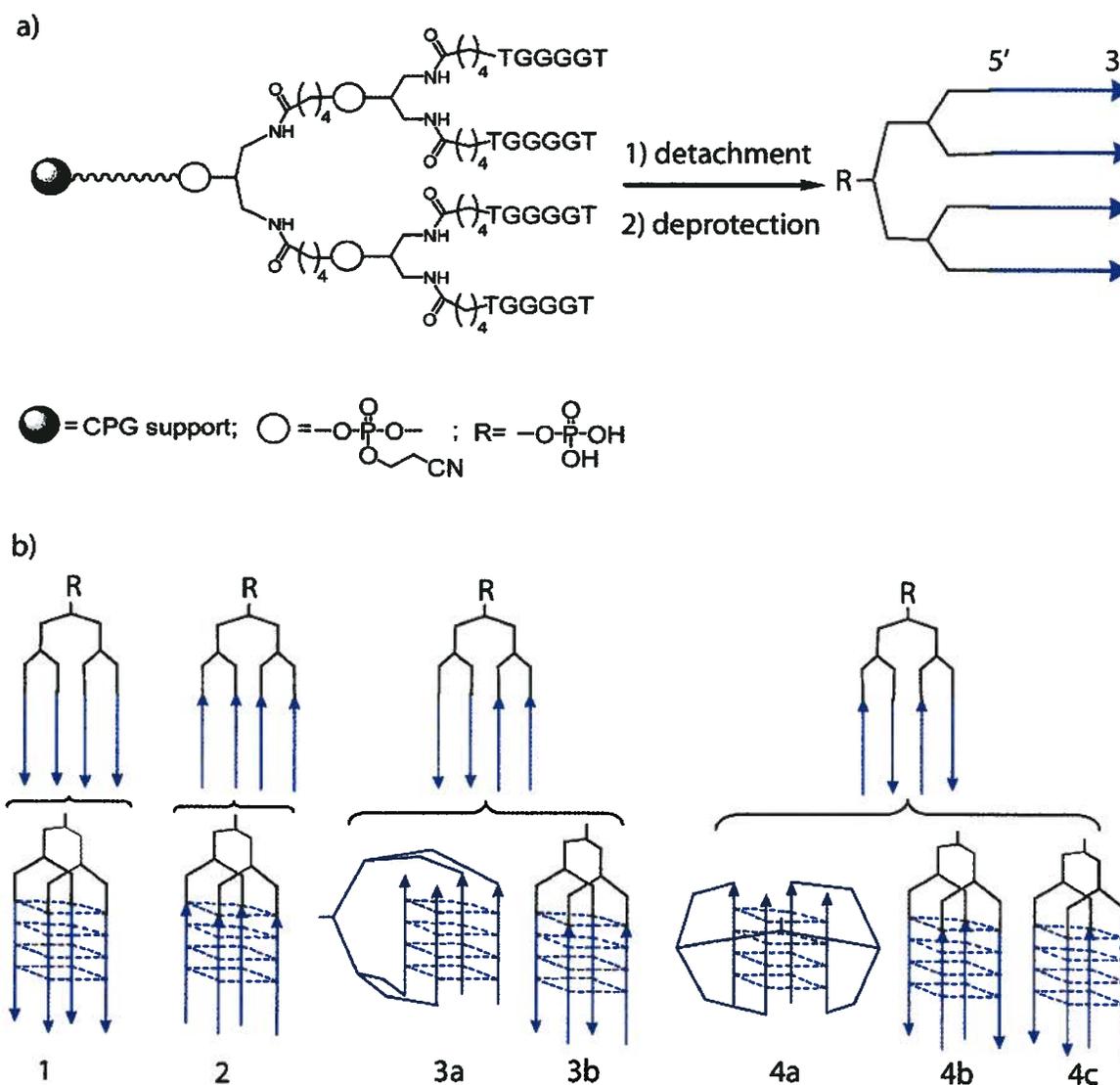
One possible area of application for PNAs is antisense therapy.<sup>128</sup> In this method the mRNA (sense) transcribed from a specific gene is targeted by a complementary strand of DNA or RNA.<sup>131</sup> Saison-Behmoaras and colleagues reported that the G-quadruplex formed from the RNA sequence of human immunodeficiency virus type 1 (HIV-1) can be specifically targeted by PNAs.<sup>132</sup> The authors have shown that the PNAs bind to the G-quadruplex and finally unfold the structure.

#### 1.4.13.2 Bunch-Oligonucleotides

Recently, Piccialli and colleagues synthesized a cluster of four oligonucleotides designed to form a unimolecular G-quadruplex in the solution.<sup>133-135</sup> The synthesis is based on the solid-phase strategy and uses a branched linker to tether the oligonucleotides (Figure 1.36a).<sup>136</sup> The linker is flexible and has been shown by molecular modeling to be able to adopt folded conformations.<sup>134</sup> The authors have synthesized several structures containing parallel and antiparallel strands and have evaluated the stabilities of the obtained G-quadruplexes using CD and <sup>1</sup>H NMR spectroscopic methods.<sup>134</sup> The structures containing parallel strands have been shown to form very stable parallel G-quadruplexes (in the presence of Na<sup>+</sup> or K<sup>+</sup> **1** and **2**, Figure 1.36b) whereas the structures containing antiparallel strands form less stable parallel G-quadruplexes due to the folding of the linker (**3a** and **4a**, Figure 1.36b). Either the antiparallel

G-quadruplexes had not been formed (**3b**, Figure 1.36b) or, if formed, were present only in minor quantities (**4b** and **4c**, Figure 1.36b). Although the bunch-oligonucleotides synthesized by Piccialli et al. are not the first unimolecular G-quadruplex reported in the literature,<sup>43</sup> their considerable stabilities make them a suitable structure for a number of investigations. They also show better kinetic and thermodynamic parameters compared to their tetramolecular analogs.

**Figure 1.36.** a) Schematic representation of the bunch-oligonucleotides (Arrows indicate 5' → 3' polarity), b) G-quadruplexes obtained from bunch-oligonucleotides in different orientations. Structure **3b** was not formed and **4b** and **4c** were obtained in minor quantities.



## **1.5 Summary and Conclusions**

This chapter was meant to give the reader an overview of the depth and diversity of G-quartet chemistry. It was shown that guanine, one of the four bases of DNA, can organize into highly-ordered structures and is a promising candidate for the formation of supramolecular assemblies. The self-assembly of guanine in lipophilic and hydrophilic systems was reviewed using relevant examples from literature, and the biological significance of nucleic acid G-quadruplexes was underscored. Moreover, it was shown that these structures are structurally tunable, and that potentially functional assemblies can be constructed from both hydrophilic and lipophilic G-quadruplexes. Chapters 2, 3, and 4 will describe the experimental work which has been carried out by the author.

## 1.6 References

- (1) Lehn, J. M. *Angew. Chem., Int. Ed.* **1990**, *29*, 1304-1319.
- (2) Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D. N.; Mammen, M.; Gordon, D. M. *Acc. Chem. Res.* **1995**, *28*, 37-44.
- (3) Schalley, C. A.; Vögtle, F.; Dötz, K. H.; Albrecht, M. *Templates in chemistry*; Berlin; [London]: Springer, c2005.
- (4) Mutter, M. *Angew. Chem., Int. Ed.* **1985**, *24*, 639-653.
- (5) Gibb, B. C.; Mezo, A. R.; Causton, A. S.; Fraser, J. R.; Tsai, F. C. S.; Sherman, J. C. *Tetrahedron* **1995**, *51*, 8719-8732.
- (6) Mezo, A. R.; Sherman, J. C. *J. Am. Chem. Soc.* **1999**, *121*, 8983-8994.
- (7) Neidle, S.; Balasubramanian, S. *Quadruplex nucleic acids*; RSC Pub.: Cambridge, 2006.
- (8) Saenger, W. *Principles of nucleic acid structure*; Springer: New York, 1984.
- (9) Neidle, S. *Nucleic acid structure and recognition*; Oxford University Press: Oxford, 2002.
- (10) Hoogsteen, K. *Acta Crystallogr.* **1963**, *16*, 907-916.
- (11) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013-2018.
- (12) Gehring, K.; Leroy, J. L.; Gueron, M. *Nature* **1993**, *363*, 561-565.
- (13) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.
- (14) Zimmerman, S. B.; Cohen, G. H.; Davies, D. R. *J. Mol. Biol.* **1975**, *92*, 181-192.
- (15) Tougaard, P.; Chantot, J. F.; Guschlbauer, W. *Biochim. Biophys. Acta* **1973**, *308*, 9-16.
- (16) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett.* **1999**, *311*, 209-214.
- (17) Pinnavaia, T. J.; Marshall, C. L.; Mettler, C. M.; Fisk, C. I.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1978**, *100*, 3625-3627.
- (18) Giorgi, T.; Grepioni, F.; Manet, I.; Mariani, P.; Masiero, S.; Mezzina, E.; Pieraccini, S.; Saturni, L.; Spada, G. P.; Gottarelli, G. *Chem.-Eur. J.* **2002**, *8*, 2143-2152.
- (19) Keniry, M. A. *Biopolymers* **2000**, *56*, 123-146.

- (20) Kilpatrick, M. W.; Torri, A.; Kang, D. S.; Engler, J. A.; Wells, R. D. *J. Biol. Chem.* **1986**, *261*, 1350-1354.
- (21) Evans, T.; Schon, E.; Goramaslak, G.; Patterson, J.; Efstratiadis, A. *Nucleic Acids Res.* **1984**, *12*, 8043-8058.
- (22) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364-366.
- (23) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593-11598.
- (24) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.
- (25) Stewart, S. A.; Weinberg, R. A. *Semin. Cancer Biol.* **2000**, *10*, 399-406.
- (26) Simonsson, T. *Biol. Chem.* **2001**, *382*, 621-628.
- (27) Davis, J. T.; Spada, G. P. *Chem. Soc. Rev.* **2007**, *36*, 296-313.
- (28) Gottarelli, G.; Masiero, S.; Spada, G. P. *J. Chem. Soc., Chem. Commun.* **1995**, 2555-2557.
- (29) Marlow, A. L.; Mezzina, E.; Spada, G. P.; Masiero, S.; Davis, J. T.; Gottarelli, G. *J. Org. Chem.* **1999**, *64*, 5116-5123.
- (30) Forman, S. L.; Fettinger, J. C.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067.
- (31) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171-182.
- (32) Kaucher, M. S.; Lam, Y. F.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *Chem.-Eur. J.* **2004**, *11*, 164-173.
- (33) Shi, X.; Fettinger, J. C.; Davis, J. T. *Angew. Chem., Int. Ed.* **2001**, *40*, 2827-2831.
- (34) Shi, X. D.; Fettinger, J. C.; Davis, J. T. *J. Am. Chem. Soc.* **2001**, *123*, 6738-6739.
- (35) Deng, J. P.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13665-13670.
- (36) Shi, X. D.; Mullaugh, K. M.; Fettinger, J. C.; Jiang, Y.; Hofstadler, S. A.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 10830-10841.
- (37) Cai, M. M.; Shi, X. D.; Sidorov, V.; Fabris, D.; Lam, Y. F.; Davis, J. T. *Tetrahedron* **2002**, *58*, 661-671.
- (38) Chantot, J. F.; Guschlbaw. *FEBS Lett.* **1969**, *4*, 173-176.

- (39) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. *Biochemistry* **1996**, *35*, 15383-15390.
- (40) Gu, J. D.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 529-532.
- (41) Włodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. *J. Phys. Chem. B* **2005**, *109*, 3594-3605.
- (42) Sen, D.; Gilbert, W. *Nature* **1990**, *344*, 410-414.
- (43) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876-880.
- (44) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189-200.
- (45) Wong, A.; Fettinger, J. C.; Forman, S. L.; Davis, J. T.; Wu, G. *J. Am. Chem. Soc.* **2002**, *124*, 742-743.
- (46) Wu, G.; Wong, A.; Gan, Z. H.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 7182-7183.
- (47) Ida, R.; Wu, G. *Chem. Commun. (Cambridge, U. K.)* **2005**, 4294-4296.
- (48) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A. *Angew. Chem., Int. Ed.* **2000**, *39*, 1300-1303.
- (49) Kotch, F. W.; Sidorov, V.; Lam, Y. F.; Kayser, K. J.; Li, H.; Kaucher, M. S.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 15140-15150.
- (50) Otero, R.; Schock, M.; Molina, L. M.; Laegsgaard, E.; Stensgaard, I.; Hammer, B.; Besenbacher, F. *Angew. Chem., Int. Ed.* **2005**, *44*, 2270-2275.
- (51) Gubala, V.; Betancourt, J. E.; Rivera, J. M. *Org. Lett.* **2004**, *6*, 4735-4738.
- (52) Liu, X. Y.; Kwan, I. C. M.; Wang, S. N.; Wu, G. *Org. Lett.* **2006**, *8*, 3685-3688.
- (53) Kaucher, M. S.; Davis, J. T. *Tetrahedron Lett.* **2006**, *47*, 6381-6384.
- (54) Sivakova, S.; Rowan, S. J. *Chem. Soc. Rev.* **2005**, *34*, 9-21.
- (55) Mezzina, E.; Mariani, P.; Itri, R.; Masiero, S.; Pieraccini, S.; Spada, G. P.; Spinozzi, F.; Davis, J. T.; Gottarelli, G. *Chem.-Eur. J.* **2001**, *7*, 388-395.
- (56) Pieraccini, S.; Gottarelli, G.; Mariani, P.; Masiero, S.; Saturni, L.; Spada, G. P. *Chirality* **2001**, *13*, 7-12.
- (57) Chandrasekhar, S. *Liquid crystals*; 2nd ed. ed.; Cambridge University Press: Cambridge, 1992.

- (58) Mariani, P.; Mazabard, C.; Garbesi, A.; Spada, G. P. *J. Am. Chem. Soc.* **1989**, *111*, 6369-6373.
- (59) Bonazzi, S.; Capobianco, M.; Demorais, M. M.; Garbesi, A.; Gottarelli, G.; Mariani, P.; Bossi, M. G. P.; Spada, G. P.; Tondelli, L. *J. Am. Chem. Soc.* **1991**, *113*, 5809-5816.
- (60) Kato, T. *Science* **2002**, *295*, 2414-2418.
- (61) Ciferri, A. *Supramolecular polymers*; 2nd ed. ed.; Taylor & Francis/CRC: New York, 2005.
- (62) Arnal-Herault, C.; Pasc, A.; Michau, M.; Cot, D.; Petit, E.; Barboiu, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 8409-8413.
- (63) Rosi, N. L.; Mirkin, C. A. *Chem. Rev.* **2005**, *105*, 1547-1562.
- (64) Daniel, M. C.; Astruc, D. *Chem. Rev.* **2004**, *104*, 293-346.
- (65) Li, Z.; Mirkin, C. A. *J. Am. Chem. Soc.* **2005**, *127*, 11568-11569.
- (66) Wu, Z. S.; Guo, M. M.; Shen, G. L.; Yu, R. Q. *Anal. Bioanal. Chem.* **2007**, *387*, 2623-2626.
- (67) Robinson, B. H.; Seeman, N. C. *Protein Eng.* **1987**, *1*, 295-300.
- (68) Dimauro, E.; Hollenberg, C. P. *Adv. Mater. (Weinheim, Ger.)* **1993**, *5*, 384-386.
- (69) Guo, X. F.; Gorodetsky, A. A.; Hone, J.; Barton, J. K.; Nuckolls, C. *Nat. Nanotech.* **2008**, *3*, 163-167.
- (70) Calzolari, A.; Di Felice, R.; Molinari, E.; Garbesi, A. *J. Phys. Chem. B* **2004**, *108*, 2509-2515.
- (71) Calzolari, A.; Di Felice, R.; Molinari, E.; Garbesi, A. *Appl. Phys. Lett.* **2002**, *80*, 3331-3333.
- (72) Delaney, S.; Barton, J. K. *Biochemistry* **2003**, *42*, 14159-14165.
- (73) de Champdore, M.; De Napoli, L.; Montesarchio, D.; Piccialli, G.; Caminal, C.; Mulazzani, Q. G.; Navacchia, M. L.; Chatgililoglu, C. *Chem. Commun. (Cambridge, U. K.)* **2004**, 1756-1757.
- (74) D'Amico, S.; Maruccio, G.; Visconti, P.; D'Amone, E.; Cingolani, R.; Rinaldi, R.; Masiero, S.; Spada, G. P.; Gottarelli, G. *Microelectron. J.* **2003**, *34*, 961-963.
- (75) Maruccio, G.; Visconti, P.; Arima, V.; D'Amico, S.; Blasco, A.; D'Amone, E.; Cingolani, R.; Rinaldi, R.; Masiero, S.; Giorgi, T.; Gottarelli, G. *Nano Lett.* **2003**, *3*, 479-483.

- (76) Gu, Q.; Cheng, C. D.; Gonela, R.; Suryanarayanan, S.; Anabathula, S.; Dai, K.; Haynie, D. T. *Nanotechnology* **2006**, *17*, R14-R25.
- (77) Sen, D.; Gilbert, W. *Biochemistry* **1992**, *31*, 65-70.
- (78) Marsh, T. C.; Henderson, E. *Biochemistry* **1994**, *33*, 10718-10724.
- (79) Marsh, T. C.; Vesenka, J.; Henderson, E. *Nucleic Acids Res.* **1995**, *23*, 696-700.
- (80) Protozanova, E.; Macgregor, R. B. *Biochemistry* **1996**, *35*, 16638-16645.
- (81) Alberti, P.; Bourdoncle, A.; Sacca, B.; Lacroix, L.; Mergny, J. L. *Org. Biomol. Chem.* **2006**, *4*, 3383-3391.
- (82) Bath, J.; Turberfield, A. J. *Nat. Nanotech.* **2007**, *2*, 275-284.
- (83) Li, J. W. J.; Tan, W. H. *Nano Lett.* **2002**, *2*, 315-318.
- (84) Alberti, P.; Mergny, J. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1569-1573.
- (85) Monchaud, D.; Yang, P.; Lacroix, L.; Teulade-Fichou, M. P.; Mergny, J. L. *Angew. Chem., Int. Ed.* **2008**, *47*, 4858-4861.
- (86) Scarso, A.; Rebek, J. J. *Top. Curr. Chem.* **2006**, *265*, 1-46.
- (87) Fasman, G. D. *Circular dichroism and the conformational analysis of biomolecules*; Plenum Press: New York, 1996.
- (88) Andrisano, V.; Gottarelli, G.; Masiero, S.; Heijne, E. H.; Pieraccini, S.; Spada, G. P. *Angew. Chem., Int. Ed.* **1999**, *38*, 2386-2388.
- (89) Dowling, V. A.; Charles, J. A. M.; Nwakpuda, E.; McGown, L. B. *Anal. Chem.* **2004**, *76*, 4558-4563.
- (90) Furlan, R. L. E.; Otto, S.; Sanders, J. K. M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 4801-4804.
- (91) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem., Int. Ed.* **2002**, *41*, 898-952.
- (92) Sreenivasachary, N.; Lehn, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5938-5943.
- (93) Sreenivasachary, N.; Lehn, J. M. *Chem.-Asian. J.* **2008**, *3*, 134-139.
- (94) Whitney, A. M.; Ladame, S.; Balasubramanian, S. *Angew. Chem., Int. Ed.* **2004**, *43*, 1143-1146.

- (95) Ladame, S.; Whitney, A. M.; Balasubramanian, S. *Angew. Chem., Int. Ed.* **2005**, *44*, 5736-5739.
- (96) Bugaut, A.; Jantos, K.; Wietor, J. L.; Rodriguez, R.; Sanders, J. K. M.; Balasubramanian, S. *Angew. Chem., Int. Ed.* **2008**, *47*, 2677-2680.
- (97) Kaucher, M. S.; Harrell, W. A.; Davis, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 38-39.
- (98) Sakai, N.; Mareda, J.; Matile, S. *Mol. Biosyst.* **2007**, *3*, 658-666.
- (99) Ghossoub, A.; Lehn, J. M. *Chem. Commun. (Cambridge, U. K.)* **2005**, 5763-5765.
- (100) Pieraccini, S.; Masiero, S.; Pandoli, O.; Samori, P.; Spada, G. P. *Org. Lett.* **2006**, *8*, 3125-3128.
- (101) Lee, J.; Kim, H. J.; Kim, J. *J. Am. Chem. Soc.* **2008**, *130*, 5010-5011.
- (102) Wu, Z. S.; Chen, C. R.; Shen, G. L.; Yu, R. Q. *Biomaterials* **2008**, *29*, 2689-2696.
- (103) Radi, A. E.; Sanchez, J. L. A.; Baldrich, E.; O'Sullivan, C. K. *J. Am. Chem. Soc.* **2006**, *128*, 117-124.
- (104) Xiao, Y.; Lubin, A. A.; Heeger, A. J.; Plaxco, K. W. *Angew. Chem., Int. Ed.* **2005**, *44*, 5456-5459.
- (105) Nagatoishi, S.; Nojima, T.; Juskowiak, B.; Takenaka, S. *Angew. Chem., Int. Ed.* **2005**, *44*, 5067-5070.
- (106) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564-566.
- (107) Lerga, T. M.; O'Sullivan, C. K. *Anal. Chim. Acta* **2008**, *610*, 105-111.
- (108) Potyrailo, R. A.; Conrad, R. C.; Ellington, A. D.; Hieftje, G. M. *Anal. Chem.* **1998**, *70*, 3419-3425.
- (109) Nutiu, R.; Li, Y. F. *J. Am. Chem. Soc.* **2003**

- (114) Ikebukuro, K.; Kiyohara, C.; Sode, K. *Anal. Lett.* **2004**, *37*, 2901-2909.
- (115) Numnuam, A.; Chumbimuni-Torres, K. Y.; Xiang, Y.; Bash, R.; Thavarungkul, P.; Kanatharana, P.; Pretsch, E.; Wang, J.; Bakker, E. *Anal. Chem.* **2008**, *80*, 707-712.
- (116) Charles, J. A. M.; McGown, L. B. *Electrophoresis* **2002**, *23*, 1599-1604.
- (117) Bourdoncle, A.; Torres, A. E.; Gosse, C.; Lacroix, L.; Vekhoff, P.; Le Saux, T.; Jullien, L.; Mergny, J. L. *J. Am. Ceram. Soc.* **2006**, *128*, 11094-11105.
- (118) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *14*, 303-308.
- (119) Baum, D. A.; Silverman, S. K. *Cell. Mol. Life Sci.* **2008**, *65*, 2156-2174.
- (120) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1995**, *2*, 655-660.
- (121) Xiao, Y.; Pavlov, V.; Niazov, T.; Dishon, A.; Kotler, M.; Willner, I. *J. Am. Chem. Soc.* **2004**, *126*, 7430-7431.
- (122) McManus, S. A.; Li, Y. F. *J. Mol. Biol.* **2008**, *375*, 960-968.
- (123) Li, Y. F.; Sen, D. *Biochemistry* **1997**, *36*, 5589-5599.
- (124) Chinnapen, D. J. F.; Sen, D. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 65-69.
- (125) Chinnapen, D. J. F.; Sen, D. *J. Mol. Biol.* **2007**, *365*, 1326-1336.
- (126) Tang, Z.; Goncalves, D. P. N.; Wieland, M.; Marx, A.; Hartig, J. S. *Chembiochem* **2008**, *9*, 1061-1064.
- (127) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D. Y.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. *Cancer Res.* **1999**, *59*, 639-644.
- (128) Nielsen, P. E. *Peptide nucleic acids: protocols and applications*; 2nd ed. / edited by Peter E. Nielsen. ed.; Horizon Bioscience: Wymondham, 2004.
- (129) Datta, B.; Schmitt, C.; Armitage, B. A. *J. Am. Chem. Soc.* **2003**, *125*, 4111-4118.
- (130) Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945.
- (131) Weiss, B.; Davidkova, G.; Zhou, L. W. *Cell. Mol. Life Sci.* **1999**, *55*, 334-358.
- (132) Boutimah-Hamoudi, F.; Leforestier, E.; Senamaud-Beaufort, C.; Nielsen, P. E.; Giovannangeli, C.; Saison-Behmoaras, T. E. *Nucleic Acids Res.* **2007**, *35*, 3907-3917.
- (133) Borbone, N.; Oliviero, G.; Amato, J.; D'Errico, S.; Galeone, A.; Piccialli, G.; Mayol, L. *Nucleosides, Nucleotides Nucleic Acids* **2007**, *26*, 1231-1236.

- (134) Oliviero, G.; Amato, J.; Borbone, N.; Galeone, A.; Petraccone, L.; Varra, M.; Piccialli, G.; Mayol, L. *Bioconjug. Chem.* **2006**, *17*, 889-898.
- (135) Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. *Tetrahedron Lett.* **2004**, *45*, 4869-4872.
- (136) Oliviero, G.; Amato, J.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. *Nucleosides, Nucleotides Nucleic Acids* **2005**, *24*, 739-741.

## **CHAPTER 2: Synthesis, Characterization, and Solution Studies of Lipophilic Template-Assembled Synthetic G-Quartets (TASQs)\***

### **2.1 Introduction**

This Chapter is organized in six sections, discussing the design, synthesis and solution behavior of the lipophilic TASQs. As defined in Chapter 1, TASQs are artificial G-quartets in which guanosine compounds are covalently connected to a cavitand template.

### **2.2 Rationale Behind Thesis**

In biological systems cation templation is the key stabilizing element of nucleic acid G-quadruplexes. H-bonding, hydrophobic interactions, and the phosphodiester backbones are other factors important for stabilizing hydrophilic G-quadruplexes.<sup>1</sup> In lipophilic systems, cation templation likewise overcomes the repulsive interactions of the carbonyl oxygens in the central core of a G-quartet.<sup>2</sup> Little attention has been paid to the role of an external backbone or templating scaffold. The present study provides a model system of how a lipophilic TASQ can be designed and synthesized with the help of an external cavitand template. The cavitand overcomes the large entropic costs of bringing guanines close together in space and eliminates the need for using high concentrations of guanosine.

---

\* "A version of this Chapter has been published. Nikan, M.; Sherman, J. C., Template-Assembled Synthetic G-Quartets (TASQs). *Angew. Chem., Int. Ed.* 2008, 47, (26), 4900-4902."

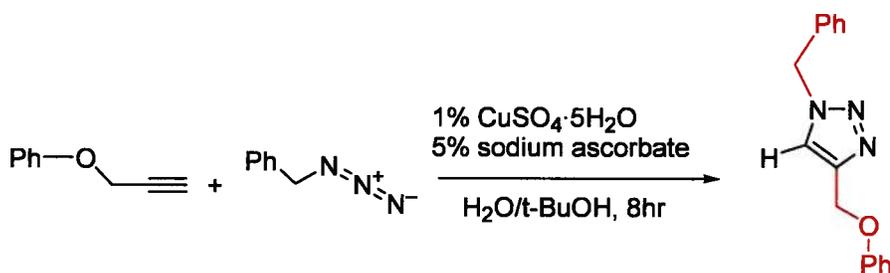
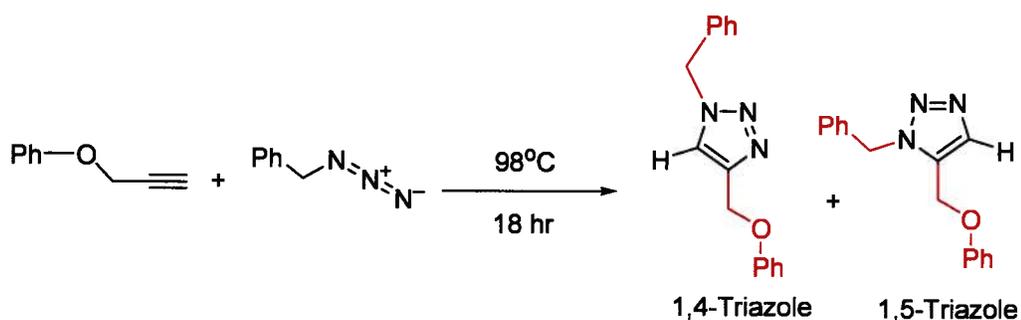
Some features of these systems worth emphasizing are: 1) The cavitand bowl is a rigid and stable compound which enhances the stability of these systems, 2) The geometry and number of functionalities on the cavitand make it possible to design and form a G-quartet with a defined topology, 3) A covalent linkage between a cavitand bowl and a guanosine compound prevents the G-quartet from converting to other structures such as G-ribbons,<sup>3</sup> 4) The template provides a simple system to probe G-quartet structures and properties, 5) The TASQs link the chemistry of G-quartets to that of cavitands, and create hybrid compounds with potential applications in biology and nanotechnology:<sup>4,5</sup> e.g. cavitand-G-quartet stationary phases for screening proteins,<sup>6</sup> and singular G-quartet baskets for encapsulation or delivery of small guests can be envisioned.

## **2.3 Results and Discussion**

### **2.3.1 TASQ via “Click” Reaction**

The synthetic strategy employed for the preparation of TASQs is based on the “click” reaction.<sup>7</sup> “Click” reaction is the name given to the copper(I)-catalyzed 1,3-dipolar cycloaddition reaction and is one of the most efficient synthetic methods for bioconjugation.<sup>8</sup> The cycloaddition reaction can be driven thermally or catalytically using a copper(I) catalyst. The thermal reaction has been known for years and is a concerted reaction that yields a mixture of two regioisomers (i.e., the 1,4 and 1,5 isomers; see Scheme 2.1).<sup>9</sup> The copper(I)-catalyzed reaction has been developed in recent years, and its stepwise mechanism results in the formation of only one regioisomer (the 1,4 isomer).<sup>10,11</sup>

**Scheme 2.1.** Thermal vs. catalytic 1,3-dipolar cycloaddition reaction

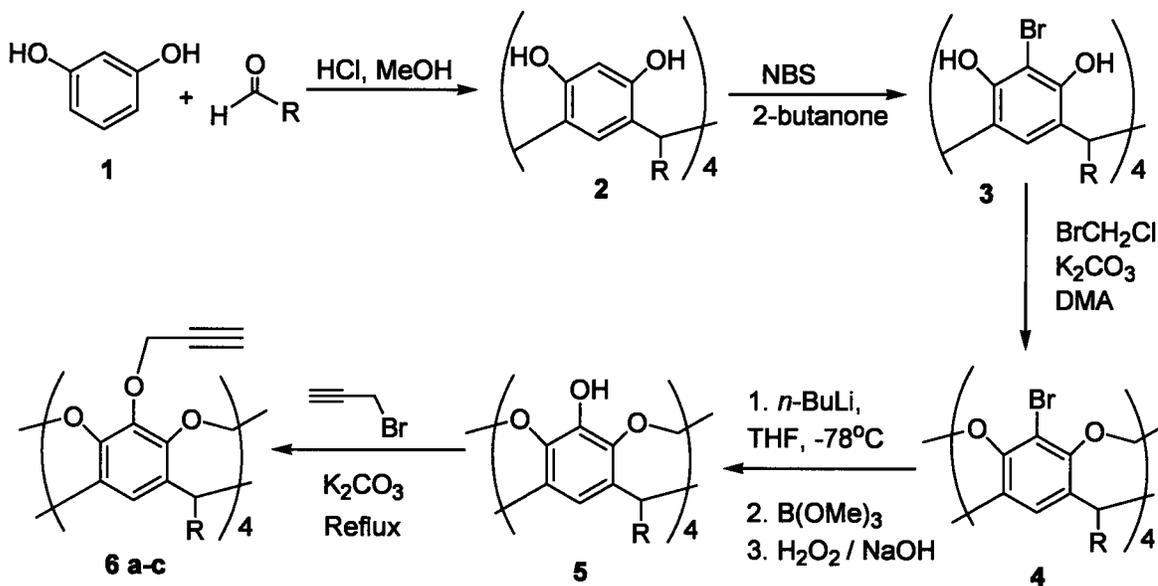


The copper(I) catalyst is typically generated *in situ* from copper(II) and an appropriate reducing agent such as sodium ascorbate.<sup>12</sup> The “click” reaction has several practical advantages: it is clean and insensitive to the presence of most functional groups. In addition, it can be performed in both aqueous and organic solution, and often gives quantitative yields. Moreover, the azide and alkynyl functionalities needed for this reaction can be readily incorporated into the reactants.<sup>10</sup> There are numerous examples in the literature in which “click” chemistry has been utilized for the decoration of different scaffolds.<sup>13-16</sup> The triazole linker obtained by this method adds yet another dimension to the chemistry of TASQs. It provides some flexibility and distance from the cavitand scaffold, and is potentially capable of  $\pi$ -stacking and H-bonding. These features might be important in the development of hydrophilic TASQs. A G-quadruplex stabilizing ligand based on triazole linker has been reported recently.<sup>17</sup>

### 2.3.2 Synthesis of Cavitands 6a-c

Cavitands **6a-c** were prepared through a five-step synthesis following literature procedures.<sup>18,19</sup> Reaction of resorcinol and aldehyde under acidic conditions afforded compound **2** after one week. Compound **2** was treated with *N*-bromosuccinimide (NBS) in 2-butanone to give bromooctol **3**. Compound **3** was then stirred with bromochloromethane and potassium carbonate in *N,N*-dimethylacetamide (DMA) to yield bromocavitand **4**. Reaction of **4** with *n*-butyllithium followed by the treatment with trimethyl borate and the oxidation with H<sub>2</sub>O<sub>2</sub> gave tetrol **5**. Subsequent treatment of **5** with propargyl bromide and potassium carbonate in acetone yielded the final products **6a-c**.

**Scheme 2.2.** The synthesis of cavitands **6a-c**.

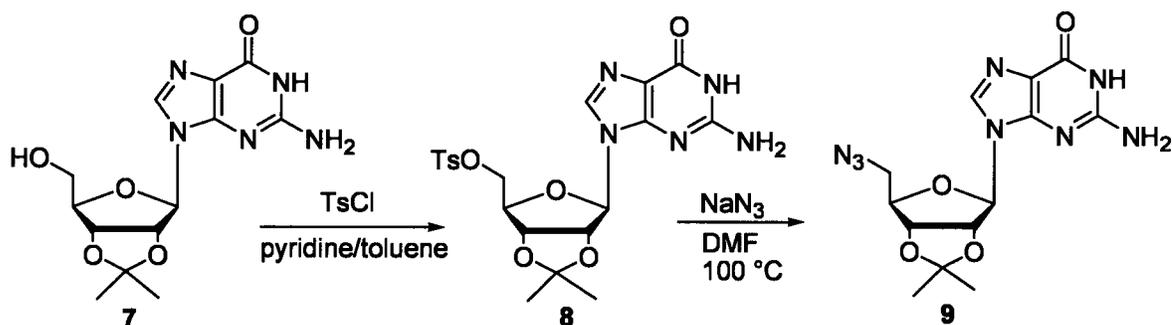


- a:**  $\text{R} = \text{CH}_3$   
**b:**  $\text{R} = \text{CH}_2\text{CH}_2 \text{ Ph}$   
**c:**  $\text{R} = \text{C}_{11}\text{H}_{23}$

### 2.3.3 Synthesis of 5'-Azido-2',3'-*O*-Isopropylidene Guanosine 9

Compound **9** was synthesized according to a known procedure.<sup>20</sup> Reaction of 2',3'-*O*-isopropylidene guanosine **7** with *p*-toluenesulfonyl chloride in pyridine/toluene gave compound **8** in 73% yield. Subsequently, reaction of **8** with sodium azide in DMF afforded 5'-azido guanosine **9** in 42% yield.

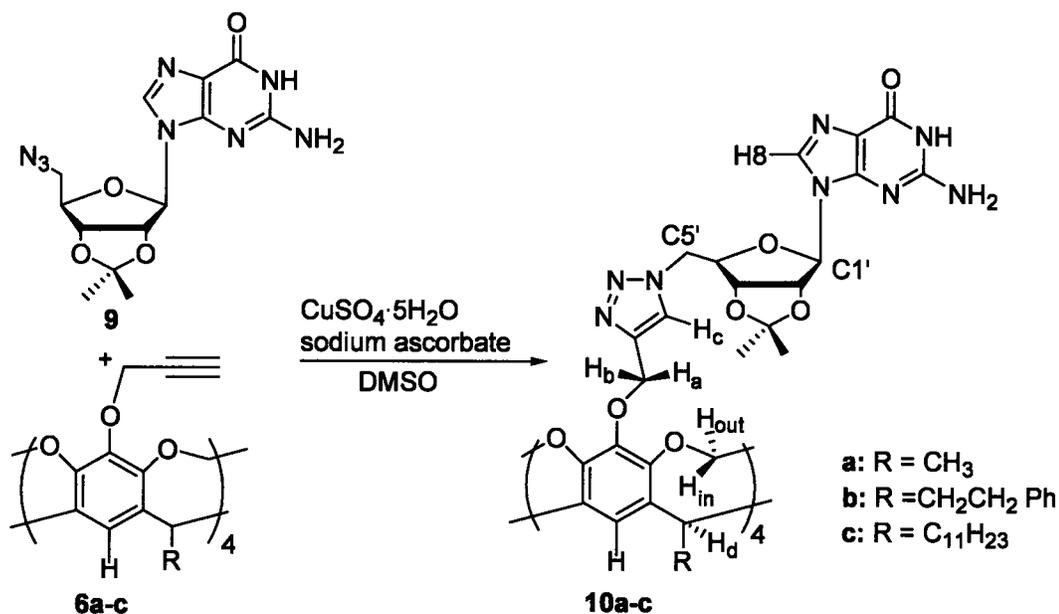
**Scheme 2.3.** The synthesis of 5'-azido guanosine **9**.



### 2.3.4 Synthesis of TASQs 10a-c

Cavitands bearing four guanosides were synthesized from cavitands **6a-c** and 5'-azido guanosine **9** through the use of the “click” reaction as shown in Scheme 2.4. The coupling reactions were performed in DMSO at 60 °C and were completed overnight. Products **10a-c** were obtained in 62-66% yield after purification on silica gel. MALDI and NMR analysis confirmed the identities of the tetrasubstituted products. Compound **10c** is freely soluble in chloroform, while compounds **10a-b** were found to be only sparingly soluble; most of the NMR experiments were performed on **10c**.

### Scheme 2.4. The synthesis of TASQs 10a-c.

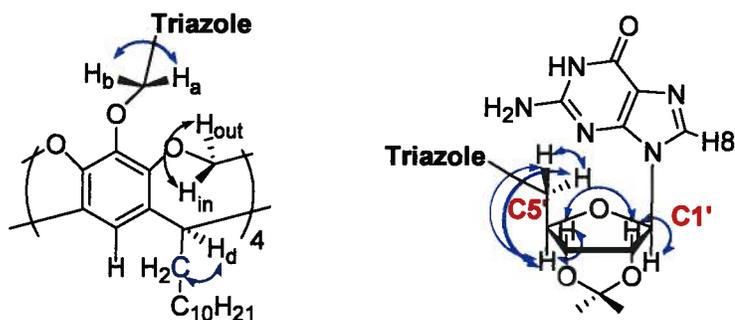


### 2.3.5 The Signal Assignment Strategy

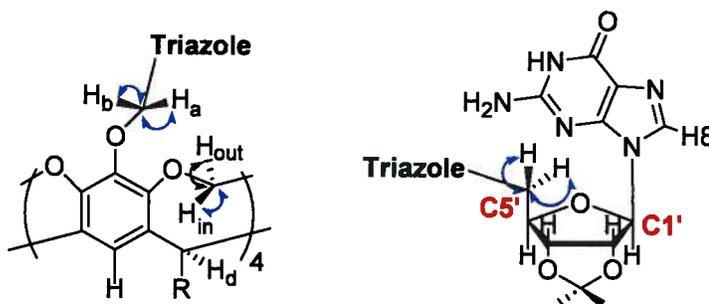
The initial NMR studies on **10c** showed marked chemical shift differences when investigated in different solvents (Figure 2.7). These findings provided the first evidence of the self-assembly of these compounds in solution and prompted us to undertake a more detailed analysis. Before providing the details of the experiments, a summary of the signal assignments is presented. Compound **10c** is a symmetrical molecule demonstrating only one set of <sup>1</sup>H NMR signals at all temperatures. The spectral assignments were achieved by a combination of <sup>1</sup>H-<sup>1</sup>H COSY, <sup>1</sup>H-<sup>13</sup>C HMQC and <sup>1</sup>H-<sup>1</sup>H NOESY. A list of important 2D NMR correlations is shown in Figure 2.1.

**Figure 2.1.** The important 2D NMR correlations observed for compound **10c** (shown broken apart for the convenience of illustrating the correlations).

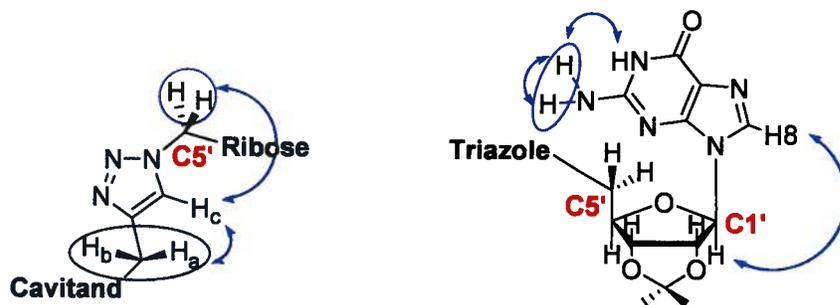
**a)**  $^1\text{H}$ - $^1\text{H}$  COSY correlations between the ribose protons,  $\text{H}_a/\text{H}_b$ ,  $\text{H}_{\text{in}}/\text{H}_{\text{out}}$ , and  $\text{H}_d/\text{cavitand feet}$ . Some correlations such as  $\text{H}1'/\text{H}2'$  were not observed due to the small coupling constants.



**b)**  $^1\text{H}$ - $^{13}\text{C}$  HMQC correlations of  $\text{H}5'/\text{H}5'b$ ,  $\text{H}_a/\text{H}_b$ , and  $\text{H}_{\text{in}}/\text{H}_{\text{out}}$  with the corresponding carbon atoms.



**c)**  $^1\text{H}$ - $^1\text{H}$  NOESY correlations between  $\text{H}1'/\text{H}8$ ,  $\text{NH}_{2a}/\text{NH}_{2b}$ ,  $\text{NH}_2/\text{NH}$ ,  $\text{H}5'/\text{H}_c$ , and  $\text{H}_{a,b}/\text{H}_c$  (an in-depth analysis of the NOE correlations illustrated here will be provided in sections 2.3.6 and 2.3.7). In addition to these important NOE correlations, the protons which are connected by scalar coupling, such as sugar protons, are generally at a close distance and show up in the NOESY spectrum.



### 2.3.5.1 Signal Assignment

This section will provide a summary of how each signal is individually assigned along with the observed chemical shifts. The signal assignments were carried out separately for the spectra recorded in CDCl<sub>3</sub> and DMSO-d<sub>6</sub> (some correlations will only be shown in one of the solvents). The results are summarized in Table 2.1. The recorded <sup>1</sup>H-<sup>1</sup>H COSY and <sup>1</sup>H-<sup>13</sup>C HMQC spectra are given in Figures 2.2, 2.3, 2.4, and 2.5.

**The ribose protons:** The ribose protons H1', H2', H3', H4', H5'a and H5'b can be unambiguously assigned based on their COSY cross-peaks. The downfield resonance of H1' is usually the starting point. Note that the COSY cross peaks for H1'/H2' and H4'/H5'a are weak

**H<sub>8</sub>:** Assigned based on its NOE cross-peak with H1' proton (Figure 2.9d).

**H<sub>c</sub>:** Assigned based on its NOE cross-peaks with H5'a/H5'b and H<sub>a</sub>/H<sub>b</sub>. These NOE correlations were also used to confirm the regiochemistry of the product (Figure 2.6b).

**H<sub>a</sub>/H<sub>b</sub>:** These geminal protons were identified by their COSY (Figures 2.2 and 2.4) and <sup>1</sup>H-<sup>13</sup>C HMQC (Figures 2.3 and 2.5) correlations and were assigned based on their NOE cross-peaks with H<sub>c</sub> (Figure 2.6b). Note that the diastereotopic pairs of H<sub>a</sub>/H<sub>b</sub> and H5'a/H5'b appear as broad signals in DMSO-d<sub>6</sub>.

**H<sub>in</sub>/H<sub>out</sub>:** Assigned based on their COSY (Figures 2.2 and 2.4) and <sup>1</sup>H-<sup>13</sup>C HMQC (Figures 2.3 and 2.5) correlations.

**H<sub>d</sub>:** Assigned based on its COSY cross-peak with the protons of cavitand feet (CH<sub>2</sub>-) (see Figure 2.2 for the cross-peak in DMSO-d<sub>6</sub>, the similar cross-peak in CDCl<sub>3</sub> has not been shown).

**ArH:** Assigned based on its known chemical shift range and NOE correlations with the protons of the cavitand feet (this NOE correlation has not been shown).

**Imino (NH) and amino (NH<sub>2</sub>) protons:** Assigned based on their known chemical shift range and NOE correlations (Figures 2.8 and 2.9c). These protons are exchangeable with D<sub>2</sub>O. The amino signals broaden up in CDCl<sub>3</sub> at RT, but become visible at low temperature (Figure 2.10).

**Table 2.1** Spectral assignments of **10c** in DMSO-d<sub>6</sub> and CDCl<sub>3</sub> at ambient temperature.

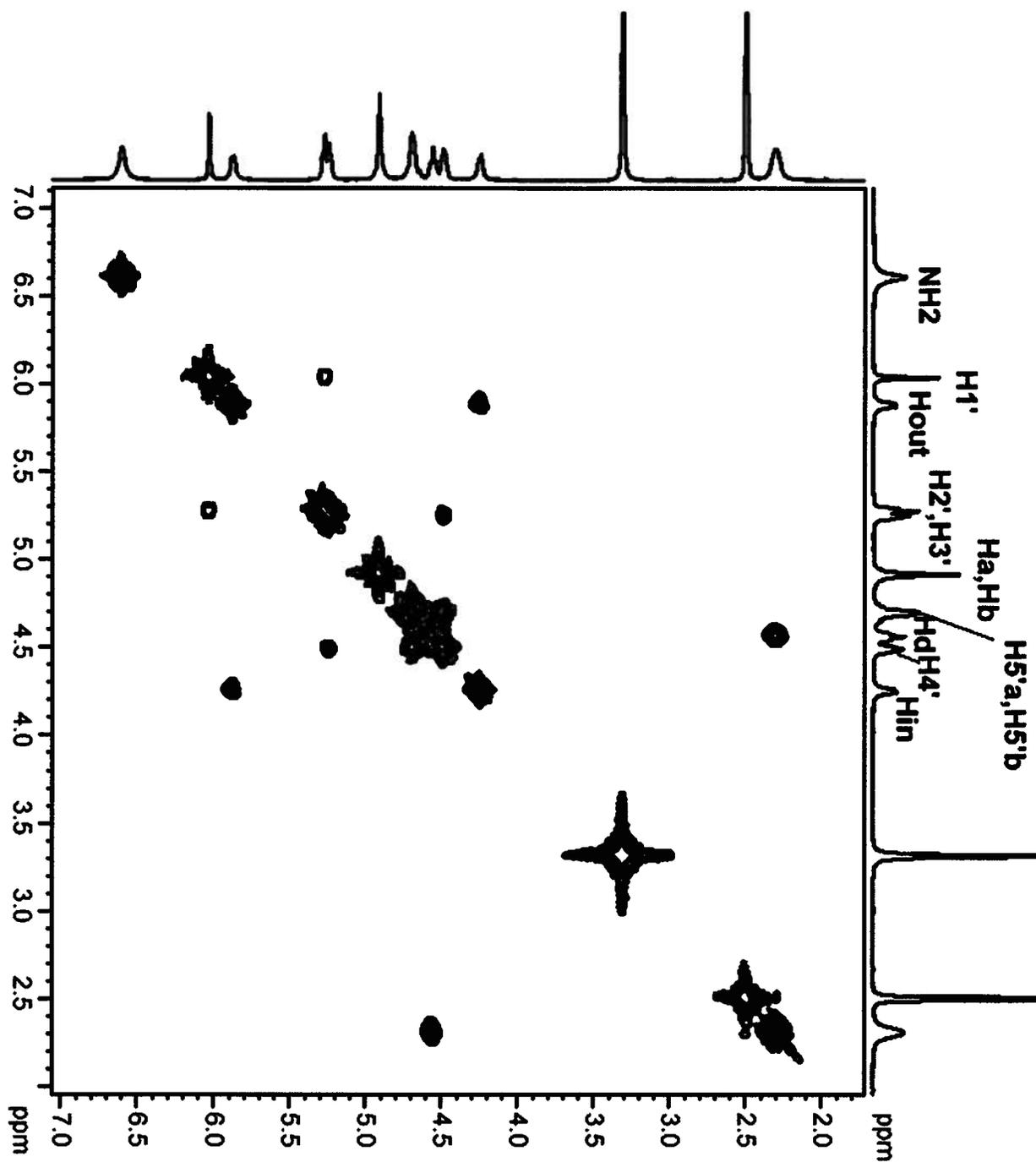
proton	$\delta_{\text{H}}$ (DMSO-d <sub>6</sub> ) <sup>a</sup>	$\delta_{\text{H}}$ (CDCl <sub>3</sub> )	<sup>1</sup> H- <sup>1</sup> H COSY <sup>b</sup>	<sup>1</sup> H- <sup>13</sup> C HMQC <sup>b</sup>
H1'	6.03	5.82	H2' *	89.17
H2'	5.27	5.14	H3', H1' *	78.44
H3'	5.24	5.13	H2', H4'	83.57
H4'	4.48	4.30	H3', H5'a*	82.90
H5'a	4.69	4.61	H5'b, H4' *	49.19
H5'b	4.69	4.96	H5'a	49.19
NH	10.78	11.63	-	-
NH <sub>2a</sub>	6.59	broad	-	-
NH <sub>2b</sub>	6.59	broad	-	-
H8	7.82	7.65	-	136.51
ArH	7.26	6.68	-	113.89
H <sub>in</sub>	4.24	4.19	H <sub>out</sub>	98.40
H <sub>out</sub>	5.88	4.50	H <sub>in</sub>	98.40
H <sub>a</sub>	4.91	4.84	H <sub>b</sub>	70.60
H <sub>b</sub>	4.91	5.70	H <sub>a</sub>	70.60
H <sub>c</sub>	8.10	7.90	-	126.13
H <sub>d</sub>	4.55	4.53	CH <sub>2</sub> (feet)	37.17

<sup>a</sup>Diastereotopic proton signals are overlapped in DMSO-d<sub>6</sub>

<sup>b</sup>2D data acquired for a 2×10<sup>-2</sup> M solution of sample in CDCl<sub>3</sub>.

\* These COSY cross peaks are very weak (due to the small coupling constants).

Figure 2.2. 400 MHz  $^1\text{H}$ - $^1\text{H}$  COSY of **10c** at 25 °C in DMSO- $d_6$ .



**Figure 2.3.** 400 MHz  $^1\text{H}$ - $^{13}\text{C}$  HMQC of **10c** at 25 °C in  $\text{DMSO-d}_6$ . The signals corresponding to diastereotopic pairs of  $\text{H5}'\text{a}/\text{H5}'\text{b}$ ,  $\text{H}_\text{a}/\text{H}_\text{b}$ , and  $\text{H}_\text{in}/\text{H}_\text{out}$  are circled.

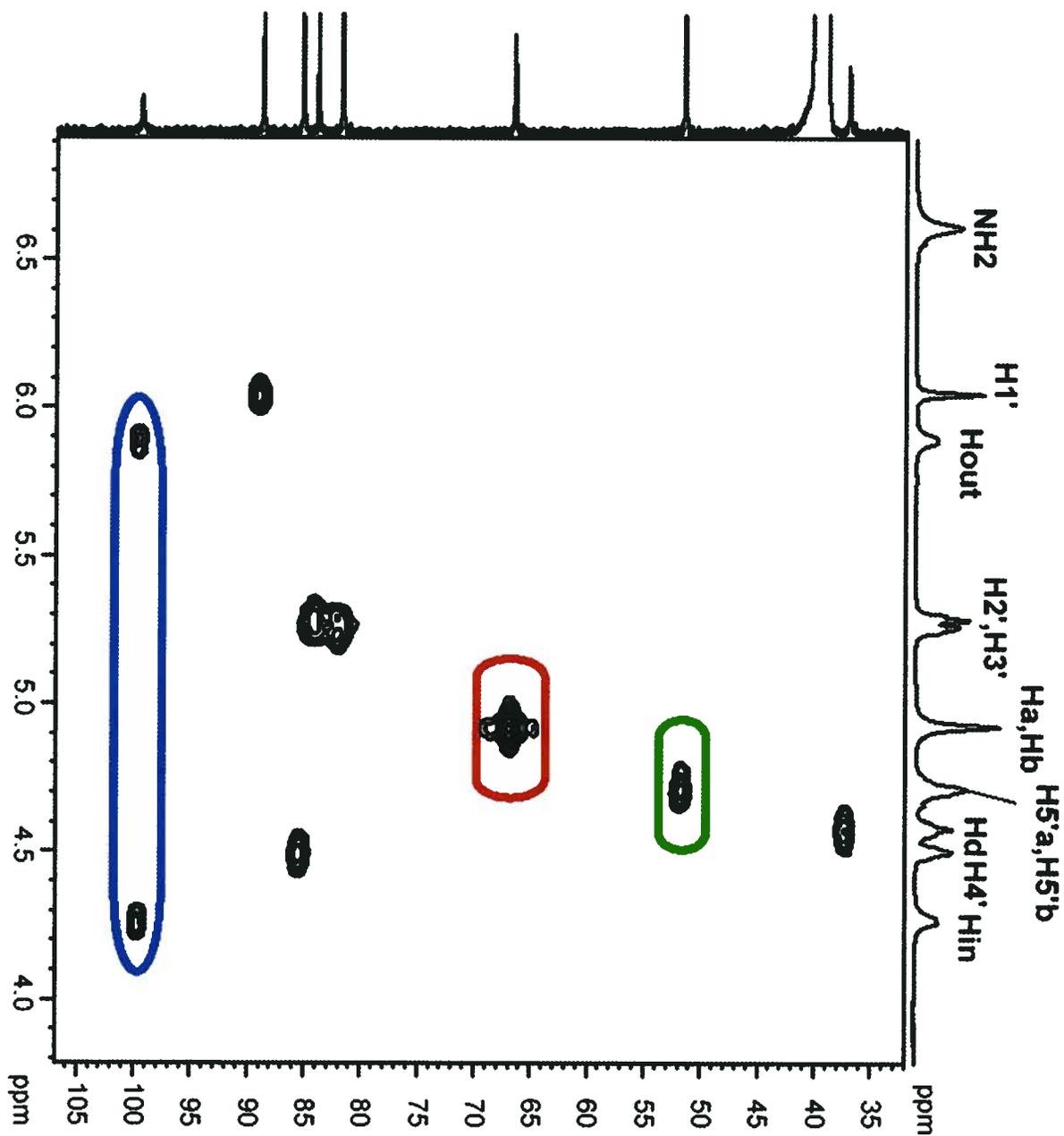
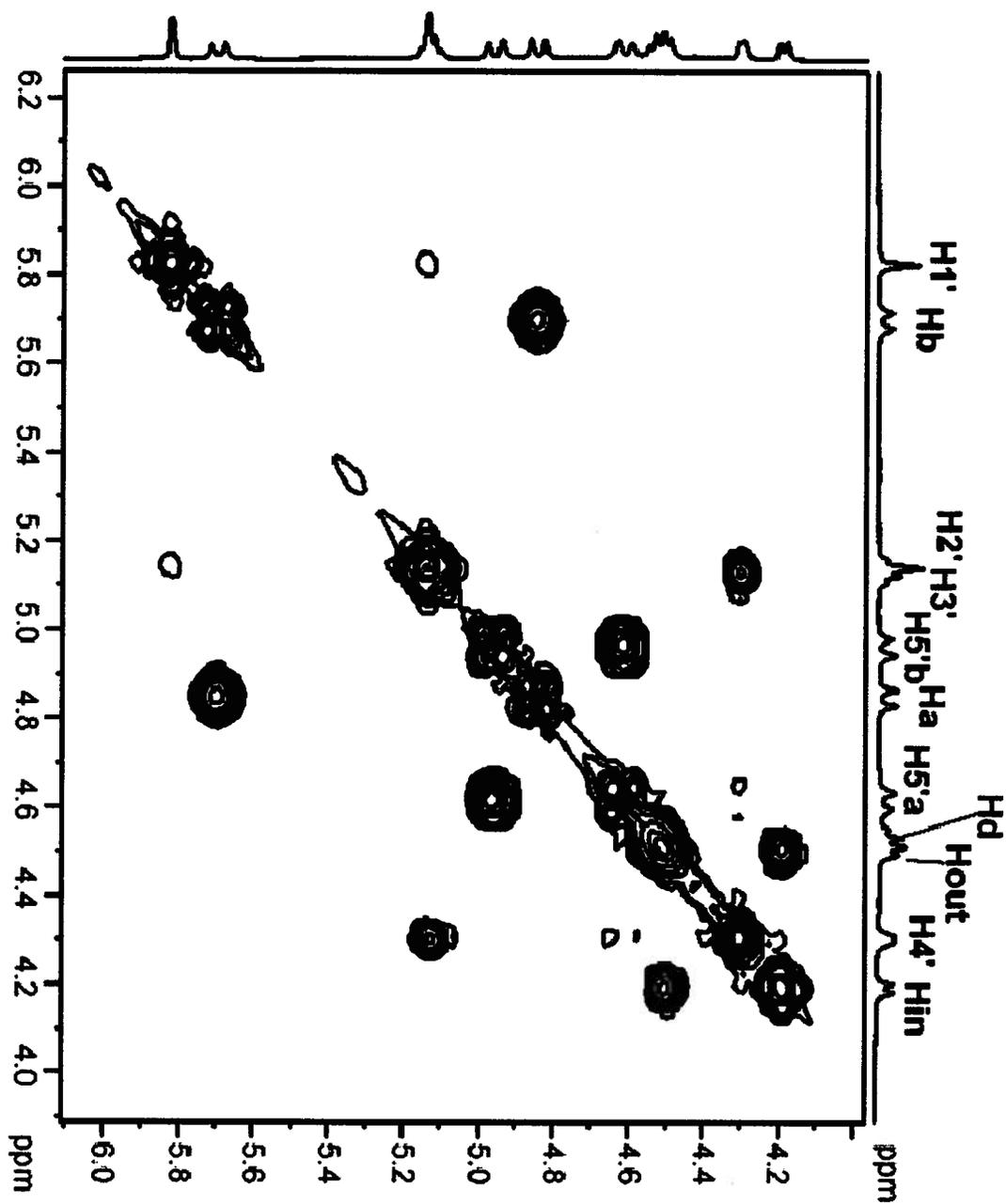
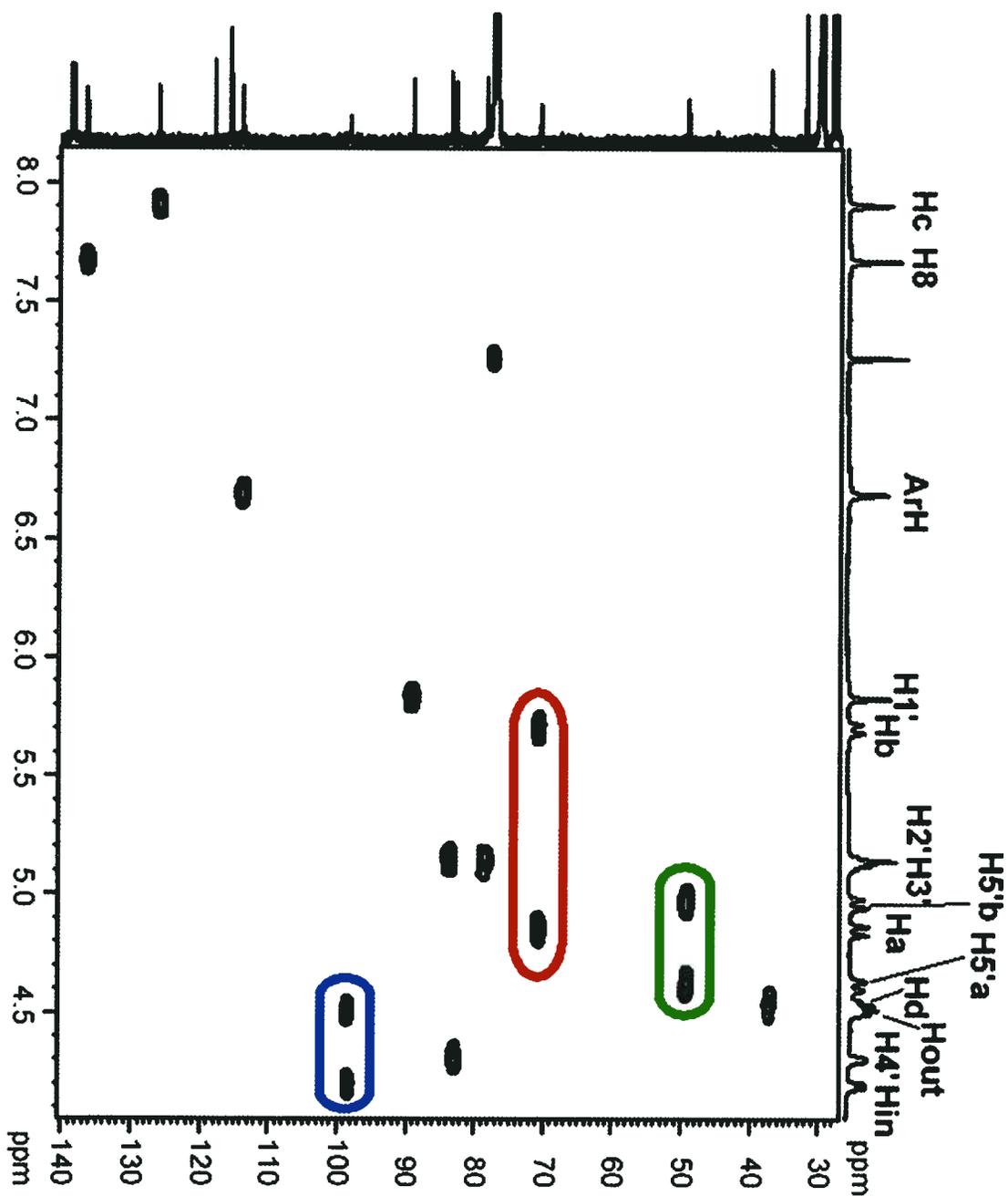


Figure 2.4. 400 MHz  $^1\text{H}$ - $^1\text{H}$  COSY of 10c at 25 °C in  $\text{CDCl}_3$ .



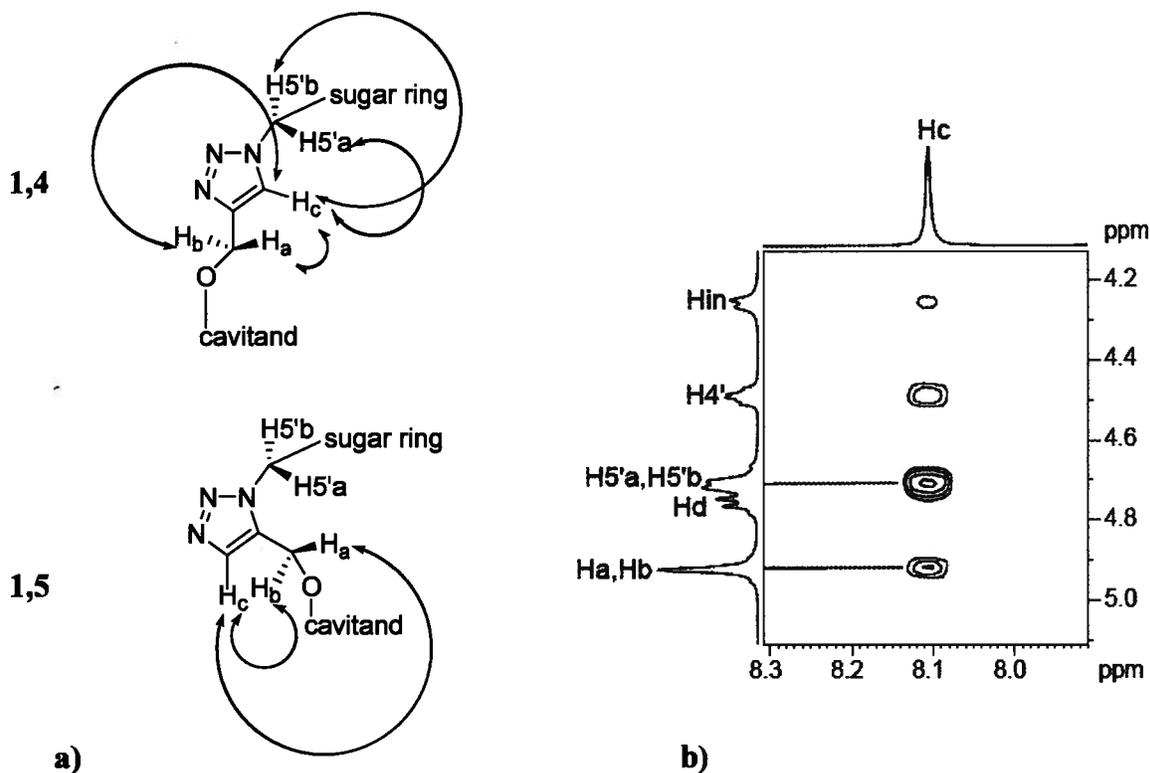
**Figure 2.5.** 400 MHz  $^1\text{H}$ - $^{13}\text{C}$  HMQC of **10c** at 25 °C in  $\text{CDCl}_3$ . The signals corresponding to diastereotopic pairs of  $\text{H5}'\text{a}/\text{H5}'\text{b}$ ,  $\text{H}_\text{a}/\text{H}_\text{b}$ , and  $\text{H}_\text{in}/\text{H}_\text{out}$  are circled.



### 2.3.6 Regiochemistry

The copper(I)-catalyzed 1,3-dipolar cycloaddition is a regioselective reaction known to yield only the 1,4-disubstituted 1,2,3-triazole.<sup>10,11</sup> The <sup>1</sup>H NMR spectrum of the product shows only one signal for the triazole proton (H<sub>c</sub>) which resonates at 8.1 ppm in DMSO. Strong NOE correlations were observed between this proton and the protons of both the linker (H<sub>a</sub>/H<sub>b</sub>) and the 5' protons of the sugar, thus confirming the structure of the 1,4-regioisomer (Figure 2.6).

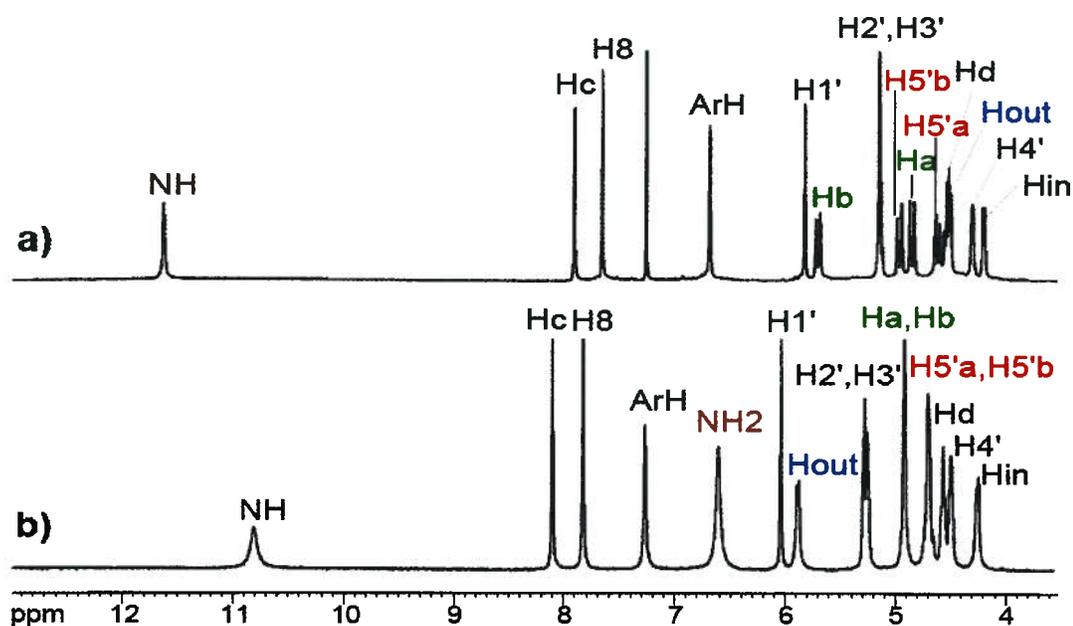
**Figure 2.6.** a) NOEs expected for 1,4 and 1,5 triazoles. b) NOEs observed for a solution of **10a** in DMSO at 400 MHz at 25 °C.



### 2.3.7 Solution Structure Determination.

As noted in the preceding sections, **10c** yields different spectra in  $\text{CDCl}_3$  and  $\text{DMSO-d}_6$  (Figure 2.7). This observation led to the suggestion that a new species is formed in  $\text{CDCl}_3$  as a result of H-bonding. A combination of 1D NMR,  $^1\text{H}$ - $^1\text{H}$  NOESY, and variable-temperature  $^1\text{H}$  NMR spectroscopic methods were used to identify the resulting structure in  $\text{CDCl}_3$ . In the following section, it will be shown that the self-assembled structure is a cation-free G-quartet.

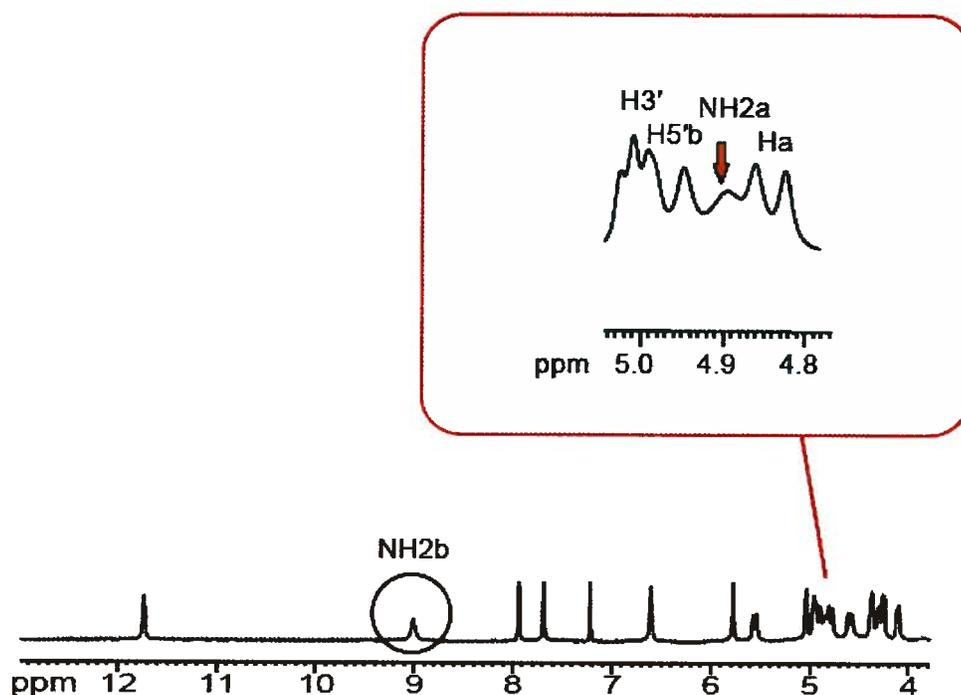
**Figure 2.7.** 400 MHz  $^1\text{H}$  NMR of **10c** at 25 °C in a)  $\text{CDCl}_3$ , and b)  $\text{DMSO-d}_6$ .



In the  $^1\text{H}$  NMR spectrum of **10c** in DMSO the amino protons ( $\text{NH}_2$ ) are equivalent and appear as a broad singlet at 6.5 ppm while the imino proton (NH) resonates at 10.7 ppm (Figure 2.7b). These chemical shift values suggest that the guanine bases are not bound in an assembly such as a G-quartet or a ribbon.<sup>21</sup>

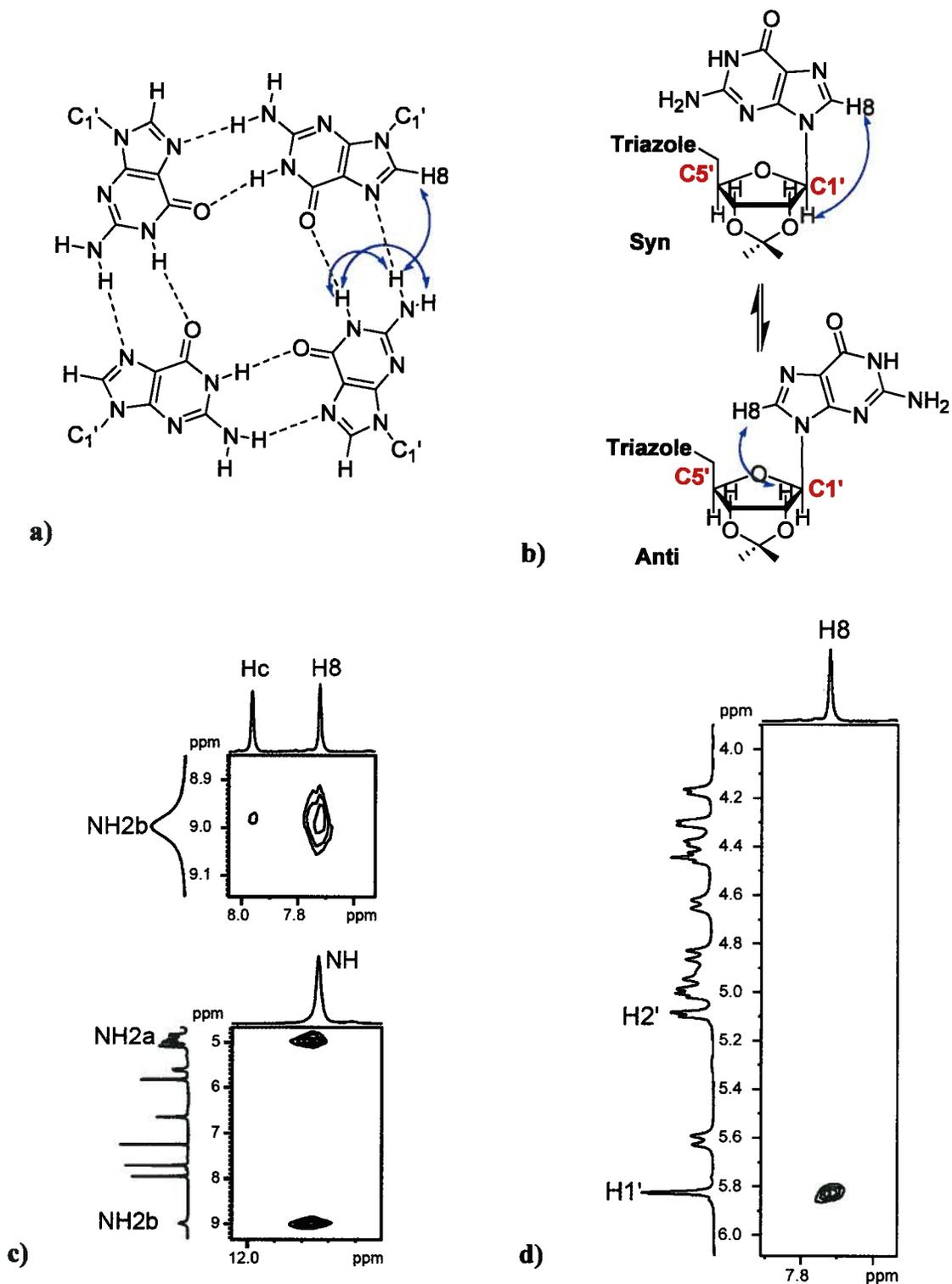
A G-quartet appears to form when **10c** is dissolved in  $\text{CDCl}_3$ , even in the absence of cations. The imino (NH) signal shifts downfield to 11.6 ppm, indicating a H-bonded system (Figure 2.7a).<sup>22,23</sup> At low temperatures, the  $\text{NH}_2$  signal appears as two distinct singlets, one at 9.0 and one at 4.9 ppm (Figure 2.8), corresponding to H-bonded and non-H-bonded protons, respectively.<sup>24</sup>

**Figure 2.8.** 400 MHz  $^1\text{H}$  NMR of **10c** at  $-40\text{ }^\circ\text{C}$  in  $\text{CDCl}_3$  indicating the H-bonded ( $\text{NH}_{2b}$ ), and the non-H-bonded ( $\text{NH}_{2a}$ ) amino signals.



More direct evidence of G-quartet formation was obtained from analysis of a 2D NOESY spectrum recorded at  $-40\text{ }^\circ\text{C}$ . A cross-peak between H8 and  $\text{NH}_{2b}$  was observed (Figure 2.9c) and has been used to authenticate a G-quartet assembly.<sup>24-26</sup> This inter-base NOE correlation plays a key role in the identification of G-quartet structures in solution (Figure 2.9a).

**Figure 2.9.** a) Inter and intra-base NOE correlations in a G-quartet, b) Intra-base NOE correlations in *syn* (strong H1'/H8 and weak H2'/H8) and *anti* (medium H1'/H8 and strong H2'/H8) conformers,<sup>24</sup> c) NOEs indicative of the formation of G-quartet, and d) NOEs indicative of the *syn* conformation at 400 MHz in CDCl<sub>3</sub> at -40 °C.



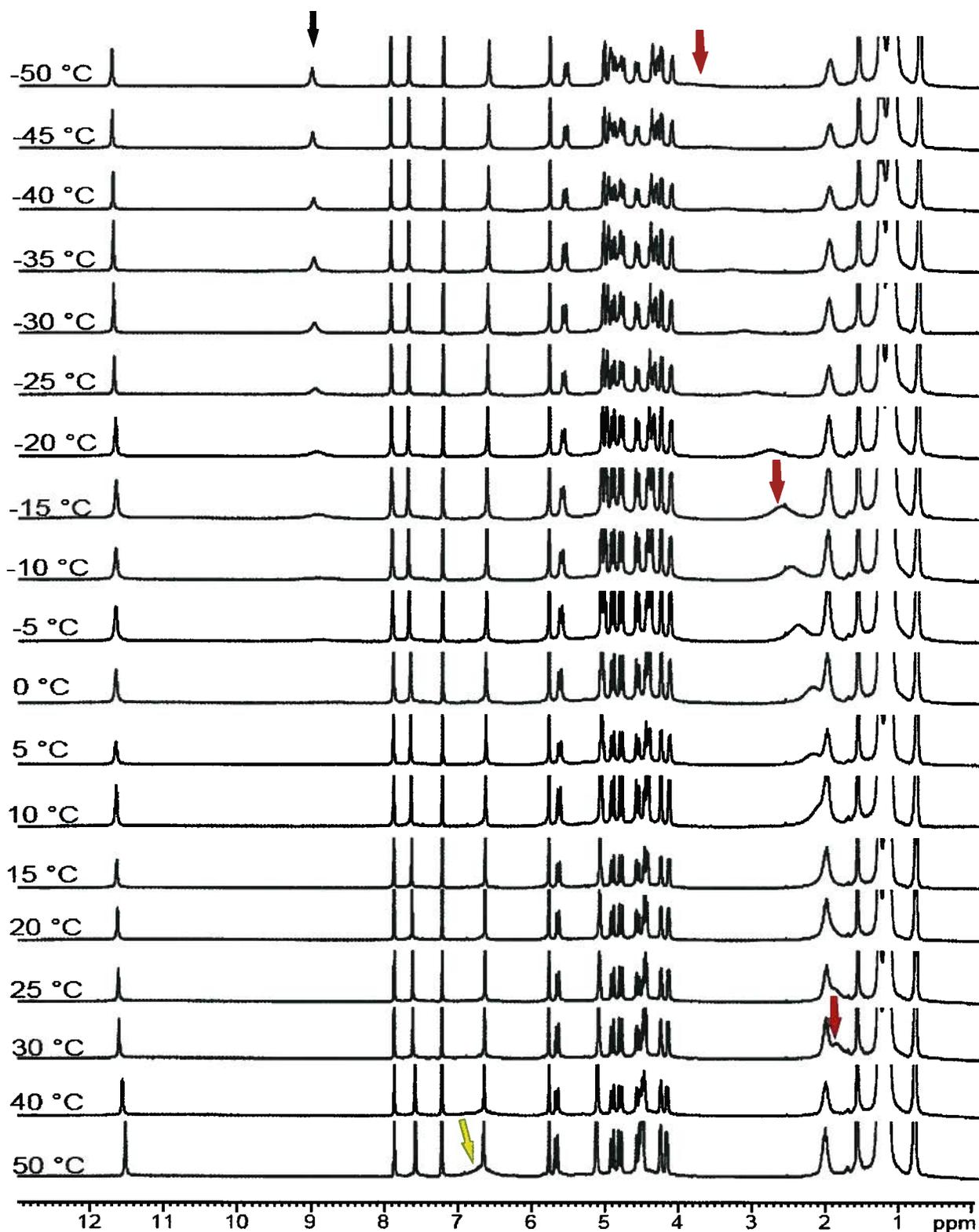
Moreover, a strong (i.e., intra-residue) NOE between H8 and H1' was observed, which is indicative of a *syn* conformation along the glycosidic bond (Figures 2.9b and 2.9d).<sup>27,28</sup> *Syn* conformations are known to prevent the formation of G-ribbons.<sup>29</sup> NOEs between amino and imino protons (Figure 2.9c) indicate Hoogsteen-paired guanines.<sup>24,30</sup> Taken together, these results suggest that **10c** spontaneously forms a G-quartet in CDCl<sub>3</sub>.

The non-exchangeable protons also exhibit changes consistent with the formation of a G-quartet. The H<sub>out</sub> proton of the cavitand undergoes a significant upfield shift in CDCl<sub>3</sub> relative to DMSO ( $\Delta\delta = -1.38$  ppm; Figure 2.7), which suggests a crowding of the upper rim of the cavitand by the aromatic guanines in CDCl<sub>3</sub>.<sup>31</sup> Diastereotopic protons H<sub>a</sub> and H<sub>b</sub>, which appear as one broad signal in DMSO, give a set of doublets in CDCl<sub>3</sub>, one of which exhibits a considerable downfield shift ( $\Delta\delta = 0.79$ ). Examination of CPK molecular models suggests that this proton (H<sub>b</sub>) is relegated to outside the anisotropic current of the aromatic rings (cavitand and guanines) upon formation of a G-quartet.

### 2.3.8 Variable-Temperature <sup>1</sup>H NMR Spectroscopic Studies

A series of variable-temperature <sup>1</sup>H NMR spectroscopic experiments were undertaken to determine the kinetic and thermodynamic stabilities of TASQs (Figure 2.10). The <sup>1</sup>H NMR spectra were recorded between -50 to 50 °C. Figure 2.10 shows that upon increasing the temperature from -50 to -10 °C, the amino protons exchange rapidly and broaden into the baseline. This broadening is the result of C-N bond rotation on the <sup>1</sup>H NMR time scale.<sup>32</sup>

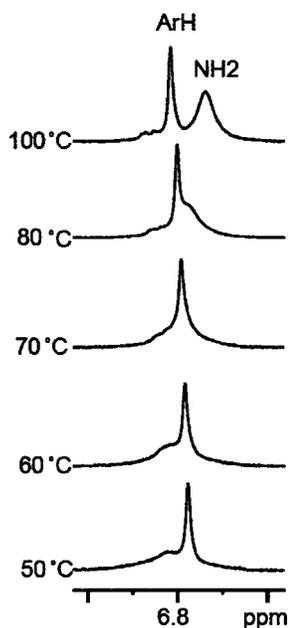
**Figure 2.10.** Variable-temperature experiments on a solution of TASQ **10c** in  $\text{CDCl}_3$  at 400 MHz. Arrows mark signals of: red: water, black: H-bonded  $\text{NH}_2$ , and yellow: the average  $\text{NH}_2$ .



In the temperature range of  $-10$  to  $30$  °C the system is still in the intermediate exchange range and the amino signals are broad. Above  $30$  °C the rotation is fast and an average signal is apparent for the  $\text{NH}_2$  protons. This kinetic stability for cation-free TASQ **10c** is comparable to the kinetic stability of some cation-bound structures.<sup>24</sup>

TASQ **10c** also appears to have considerable thermodynamic stability (Table 2.2). There is only a small change in the chemical shift of the imino (NH) signal over a  $100$  °C temperature range ( $\Delta\delta = -0.2$ ). This indicates that the H-bonding remains largely intact even at  $50$  °C in  $\text{CDCl}_3$ . The amino signal changes are harder to follow as the coalesced signal only becomes observable at temperatures close to the boiling point of  $\text{CDCl}_3$ . A variable-temperature  $^1\text{H}$  NMR experiment in  $\text{CDCl}_2\text{-CDCl}_2$  showed that the amino signal sharpens and shifts upfield ( $\Delta\delta = 0.15$  ppm) at higher temperatures ( $50$ - $100$  °C). This observation suggests that the amino proton H-bond is not completely disrupted at  $50$  °C (Figure 2.11).

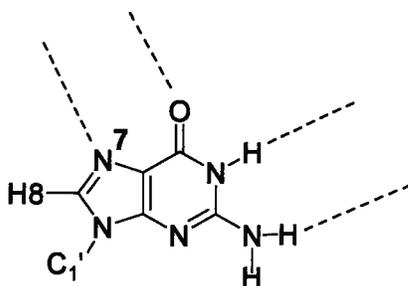
**Figure 2.11.** Changes in the chemical shift of coalesced  $\text{NH}_2$  signal of **10c** in  $\text{CDCl}_2\text{-CDCl}_2$  from  $50$ - $100$ °C.



The chemical shift of H<sub>out</sub> also manifests modest temperature-sensitivity ( $\Delta\delta = 0.16$  ppm,  $-50$  to  $+50$  °C in CDCl<sub>3</sub>). Considering that H<sub>out</sub> has a solvent-dependent  $\Delta\delta$  (DMSO-CDCl<sub>3</sub>) of 1.38 ppm, the 0.16 ppm  $\Delta\delta$  observed here with a 100 °C temperature change in CDCl<sub>3</sub> appears to represent a small degree of denaturation.

Small chemical shift changes were also observed for H8 ( $\delta_{+50^\circ\text{C}} - \delta_{-50^\circ\text{C}} = -0.11$  ppm) and H<sub>b</sub> ( $\delta_{+50^\circ\text{C}} - \delta_{-50^\circ\text{C}} = +0.11$  ppm) over this 100 °C temperature range in CDCl<sub>3</sub>. The upfield shift of H8 could be attributed to a minor change in the electron density near this proton caused by the adjacent nitrogen (N7). The electron withdrawing effect of N7 on H8 decreases as this nucleophilic site becomes less involved in H-bonding (Figure 2.12). The H<sub>b</sub> proton shift is in a conformationally more complex situation. However this chemical shift change is still small compared to the shift observed for this proton by the change of the solvents. Although the extent of H-bonding decreases at higher temperatures, chemical shift values for the exchangeable and non-exchangeable protons affected by the H-bonding suggest that the TASQ 10c retains considerable G-quartet structure even at elevated temperatures (Table 2.2). It should be noted that only one set of signals was observed during this experiment, indicating that only a time-average of all species is observed by <sup>1</sup>H NMR. A final note: our experiment showed no concentration dependence for the observed NMR spectra in the range of 0.2-10 mM.

**Figure 2.12.** The donor and acceptor faces of guanine in a G-quartet unit.



**Table 2.2** Table of chemical shifts for a solution of **10c** in CDCl<sub>3</sub> at four different temperatures.

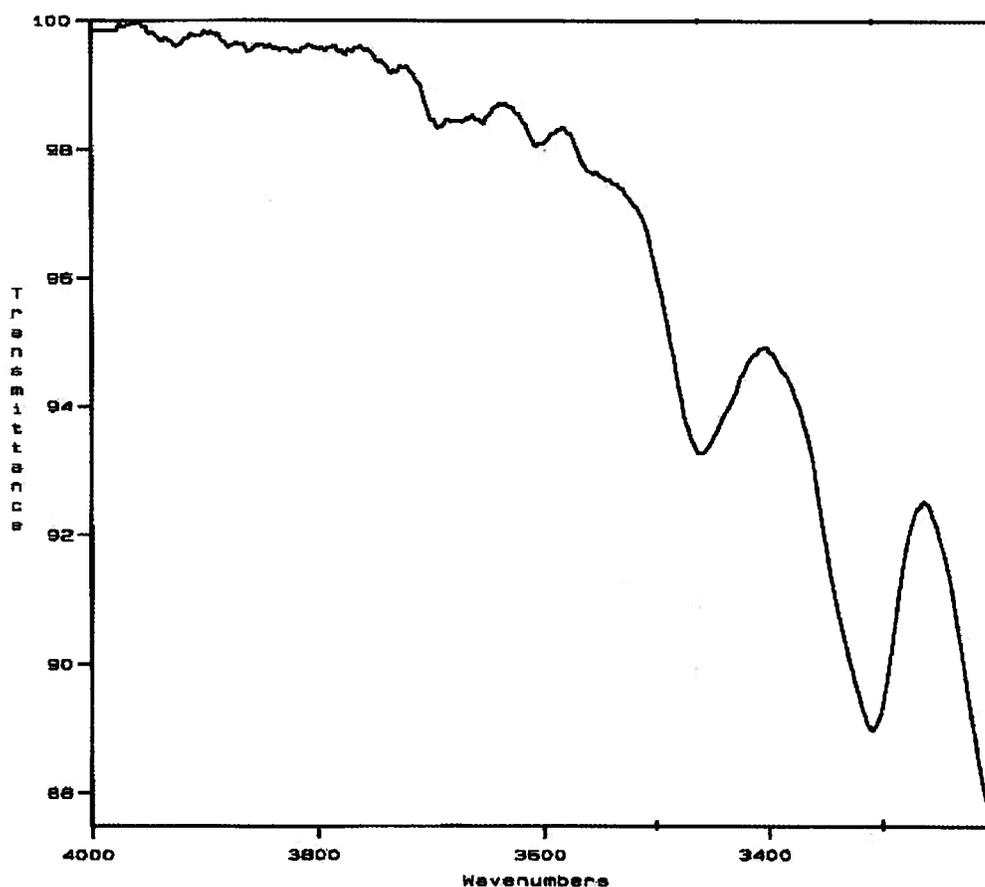
proton	$\delta_{\text{H}}$ (-50 °C)	$\delta_{\text{H}}$ (0 °C)	$\delta_{\text{H}}$ (25 °C)	$\delta_{\text{H}}$ (50 °C)
H1'	5.82	5.82	5.82	5.82
H2'	5.09	5.13	5.14	5.17
H3'	5.00	5.08	5.13	5.17
H4'	4.31	4.30	4.30	4.30
H5'a	4.65	4.62	4.61	4.60
H5'b	4.97	4.96	4.96	4.96
NH	11.74	11.67	11.63	11.54
NH <sub>2a</sub>	4.90	broad	broad	broad
NH <sub>2b</sub>	9.03	broad	broad	broad
H8	7.72	7.69	7.65	7.61
ArH	6.65	6.67	6.68	6.69
H <sub>in</sub>	4.17	4.18	4.19	4.22
H <sub>out</sub>	4.36	4.46	4.50	4.53
H <sub>a</sub>	4.84	4.84	4.84	4.85
H <sub>b</sub>	5.60	5.68	5.70	5.71
H <sub>c</sub>	7.97	7.92	7.90	7.90
H <sub>d</sub>	4.43	4.50	4.53	4.53

### 2.3.9 Infrared Spectroscopic Studies

In addition to the results obtained by variable-temperature <sup>1</sup>H NMR spectroscopy, the information regarding the H-bonding of guanines at RT can be gathered by examining the IR spectrum. Guanine bases exhibit distinct absorption bands due to symmetric and asymmetric N-

H stretching vibrations.<sup>33</sup> Depending on whether the amino (NH<sub>2</sub>) or imino (NH) groups are free or H-bonded, these absorption bands appear at different regions. The infrared spectrum of TASQ 10c in CDCl<sub>3</sub> shows strong association bands for the amino (NH<sub>2</sub>) and imino (NH) groups which suggest that they are both involved in H-bonding at RT (Figure 2.13).

**Figure 2.13.** The amino (NH<sub>2</sub>) and imino (NH) stretching region (3200-3600 cm<sup>-1</sup>) of the infrared spectrum of a 3×10<sup>-2</sup> M solution of TASQ 10c in CDCl<sub>3</sub> at RT. The association bands at 3310 and 3470 cm<sup>-1</sup> are due to H-bonding.



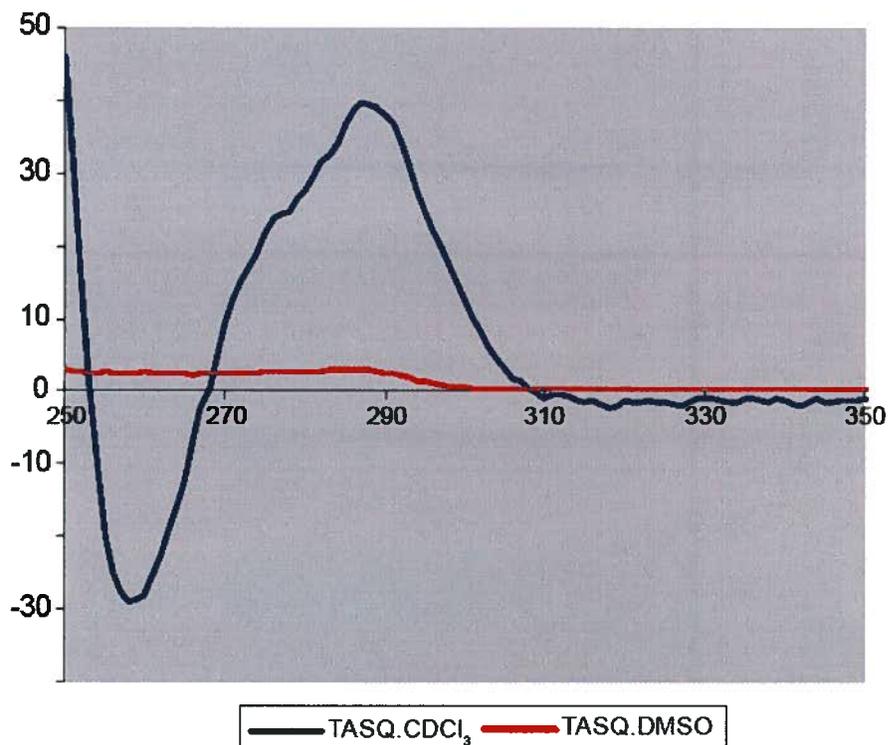
The infrared spectrum of guanine bases in CDCl<sub>3</sub> have been investigated by different research groups in the past.<sup>22,33</sup> They have assigned the absorption bands at 3515 and 3407 cm<sup>-1</sup>

as indicative of free amino ( $\text{NH}_2$ ), the band at  $3385\text{ cm}^{-1}$  as indicative of free imino ( $\text{NH}$ ), and the bands at  $3480$ ,  $3355$ ,  $3305$ , and  $3230\text{ cm}^{-1}$  as association bands due to the H-bonding of guanines.<sup>33</sup> As it is shown in Figure 2.13, the absorption bands corresponding to free  $\text{NH}$  and  $\text{NH}_2$  groups are not seen in the IR spectrum of TASQ **10c** in solution and the spectrum is dominated by two strong association bands at  $3310$  and  $3470\text{ cm}^{-1}$ . Although the most compelling evidence for the self-assembly and thermodynamic stability of TASQ **10c** comes from the NMR analysis, the IR spectrum confirms that the guanine bases are H-bonded at RT.

### 2.3.10 CD Spectroscopic Studies.

Furanose sugars are chiral and induce a small CD band in the electronic transition of a single nucleotide. When guanines are arranged in a G-quadruplex this CD band is intense and gives valuable information about the structures in solution.<sup>34</sup> The CD spectrum of TASQ **10c** in chloroform exhibits a maximum at ca.  $288\text{ nm}$  and a minimum at ca.  $259\text{ nm}$  (Figure 2.14). However, this CD band cannot be attributed only to the H-bonding of guanines.<sup>35,36</sup> Cavitand and triazole are two other chromophores present which absorb in the same region and contribute to the observed CD bands. As shown in Figure 2.14, the CD bands that TASQ **10c** manifests in chloroform is absent in DMSO, which is consistent with the  $^1\text{H}$  NMR data that suggests the loss of original structure in the polar DMSO solvent (Figure 2.14).

**Figure 2.14.** CD spectra of a 0.2 mM solution of TASQ **10c** in chloroform and DMSO.



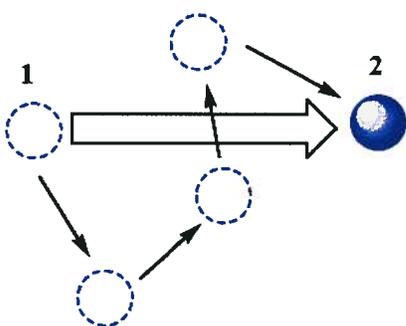
### 2.3.11 Diffusion NMR Studies

Guanosine compounds are able to self-associate and form higher order structures in the solution.<sup>2</sup> These structures are generally symmetrical and may exhibit identical NMR spectra (e.g. tetramer vs. octamer). The results presented so far rest on the assumption that TASQ **10c** is a well-defined unimolecular entity in the solution. This assumption needed to be verified. The following sections will detail the size determination studies of TASQ **10c** using the pulse field-gradient (PFG) NMR spectroscopic method. In the next few pages a brief description of PFG NMR spectroscopy is provided followed by the characterization data. The obtained data is consistent with the presence of an isolated G-quartet in CDCl<sub>3</sub>.

### 2.3.11.1 Translational Self-Diffusion

Translational diffusion is the random thermal motion of molecules (or ions) in a specific period of time (Figure 2.15).<sup>37</sup>

**Figure 2.15.** The translational diffusion of a molecule from point 1 to point 2, which is independent of its path.



The diffusion coefficient ( $D$ ) of an unhydrated spherical molecule is related to its hydrodynamic frictional coefficient and hydrodynamic radius by the Stokes-Einstein equation (Equation 1).<sup>37,38</sup>

$$D = \frac{k_b \cdot T}{f} = \frac{k_b \cdot T}{6 \cdot \pi \cdot \eta \cdot R_h} \quad (1)$$

Where  $D$  refers to the diffusion coefficient ( $\text{cm}^2 \cdot \text{s}^{-1}$ ),  $f$  is the hydrodynamic frictional coefficient,  $T$  is the temperature in Kelvin,  $k_b$  is the Boltzmann constant ( $\text{g} \cdot \text{cm}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$ ),  $\eta$  is the viscosity in Poise ( $\text{g} \cdot \text{cm}^{-1} \cdot \text{s}^{-1}$ ), and  $R_h$  is the hydrodynamic or Stokes radius (cm). The hydrodynamic frictional coefficient of non-spherical molecules is derived from other equations, and will generally be larger than that of spherical molecules with the same volume. In other

words, non-spherical molecules in general have larger surface to volume ratios. In principle, it is possible to calculate the hydrodynamic frictional coefficient of any molecule using the computational or analytical methods.<sup>37,39</sup> There is one more factor which has to be taken into account in calculating the frictional coefficients. This factor, which is more important for biological molecules, is solvation. For fully solvated molecules such as proteins, the hydrodynamic radius ( $R_h$ ) is smaller than the effective radius of the hydrated molecules. The solvation-factor can be ignored for apolar solvents.<sup>40,41</sup>

The hydrodynamic radius ( $R_h$ ) of a spherical molecule is related to its shape, molecular weight, and partial specific volume by Equation 2.

$$V = \frac{4}{3} \pi \cdot R_h^3 = \frac{M \cdot \bar{v}}{N_A} \quad (2)$$

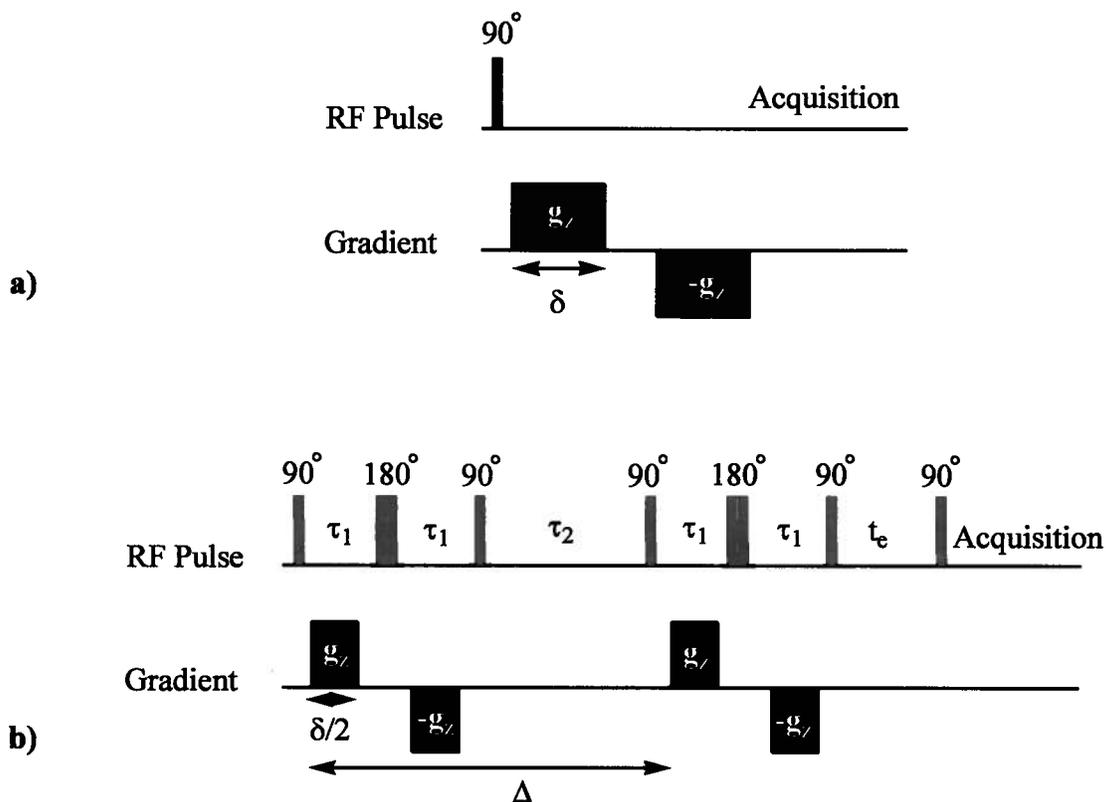
Where  $V$  is the hydrodynamic volume,  $M$  is the molecular weight,  $N_A$  is Avogadro's number, and  $\bar{v}$  is the partial specific volume. The partial specific volume is defined as the volume of solution (in mL) occupied per gram of unhydrated solute, and is experimentally determinable.<sup>42</sup>

### **2.3.11.2 The Theory of Pulse Field-Gradient (PFG) NMR Spectroscopy**

There are different methods for measuring a diffusion coefficient. Among those, NMR methods are the most popular due to their ease, sensitivity, and nondestructive nature.<sup>43</sup> It has been first demonstrated by Stejskal and Tanner that translational diffusion of molecules can be measured by NMR spectroscopy using gradient pulses.<sup>44</sup> There are currently various methods

for conducting the experiments.<sup>43</sup> One of them involves using the bipolar pulse longitudinal eddy current delay (BPLED) sequence which has been utilized in the present study (Figure 2.16b). The simple gradient pulse sequence of (a) in the same figure is provided to facilitate the explanation of each step in this process.<sup>45</sup>

**Figure 2.16.** a) A simple gradient pulse sequence, b) The BPLED sequence. In this sequence, short gradients of opposite polarity are used, which are separated by a  $180^\circ$  pulse. This combination reduces the eddy current effects (induced by gradient pulses), and improves the line shape.



A magnetic field gradient is a variation in the magnetic field along a specific set of coordinate axes. Thus, a one-dimensional gradient is a variation in magnetic field in one direction (x, y or z).<sup>46</sup> Most modern NMR spectrometers have gradient systems which are

capable of generating magnetic field gradient pulses along the z-direction.<sup>47</sup> In the absence of an external gradient, the nuclear spins are in a homogeneous magnetic field ( $B_0$ ), and precess with their Larmor frequencies. The initial  $90^\circ$  pulse (Figure 2.16a) rotates the magnetization from the z axis to the xy plane. Then the first external gradient is applied (Figure 2.16a), and each spin undergoes a phase shift due to the applied gradient. This stage is called encoding or dephasing. After waiting for a short period of time, the second gradient which has equal magnitude and duration, and opposite sign is applied (Figure 2.16a). What happens next depends on this waiting period. When the waiting period is too short the nucleus will not have a chance to change position. Therefore, the decoding or rephrasing stage will result in refocusing of magnetization, and the net phase change will be zero. However, if the spin diffuses during this time period, the net phase change will not be zero. The addition of the phases over the whole sample volume will lead to partial decrease of magnetization and therefore attenuation of the measured signal.<sup>45,46</sup> Stejskal and Tanner showed that signal attenuation is related to the diffusion coefficient by Equation 3.<sup>44</sup>

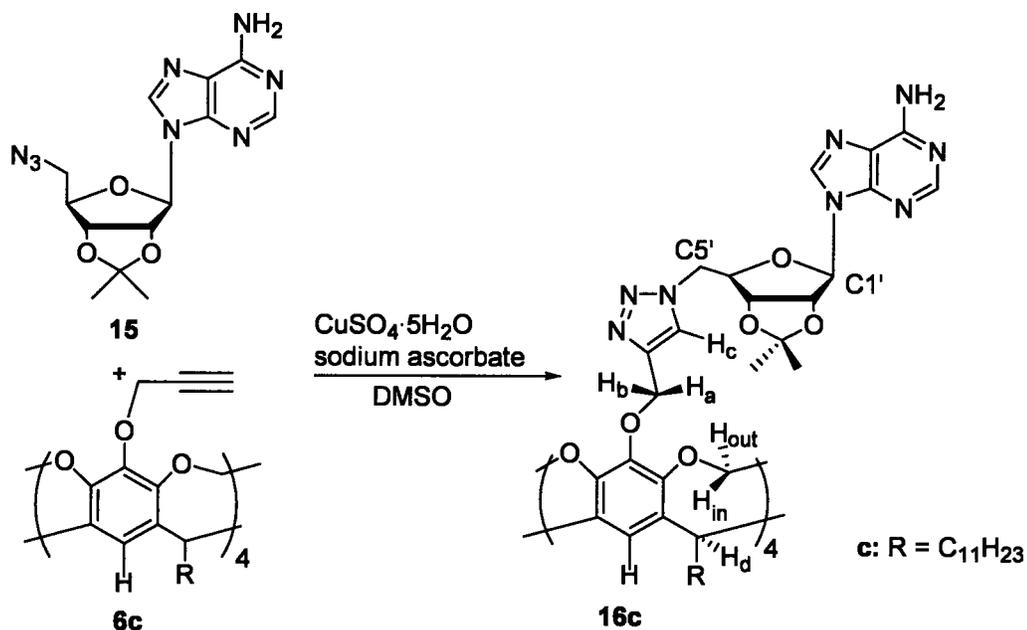
$$\ln \frac{I}{I_0} = -D(2\pi\gamma g\delta)^2 \left( \Delta - \frac{\delta}{3} \right) = -Db \quad (3)$$

As shown by Equation 3, the diffusion coefficients can be calculated by plotting the natural logarithm of the signal intensity versus the b-value. The slope of this line equals  $-D$ . The b value is the diffusion weighting factor which is defined as  $b = (2\pi\gamma g\delta)^2(\Delta - \delta/3)$  s/m<sup>2</sup> (in which  $\gamma$  is the gyromagnetic ratio,  $\delta$  is the gradient pulse duration, and  $g$  and  $\Delta$  are the gradient strength and the time between the start of the first and the second gradient pulses).

### 2.3.11.3 The Size Determination of TASQ 10c by PFG NMR Spectroscopy

Pulse field-gradient NMR spectroscopy has gained increasing importance in the field of supramolecular chemistry in recent years.<sup>43,48</sup> This technique has been successfully applied to the size determination of a variety of self-assembled structures such as resorcinarenes,<sup>49,50</sup> (iso)guanosines<sup>51,52</sup> and rosettes.<sup>53</sup> For more accurate measurements, it is often necessary to use an internal standard with a similar structure. In the past, adenine compounds have been used as internal standards for diffusion study of guanine compounds.<sup>51,54</sup> Adenine is believed to show few interactions with itself or with guanine, and has a similar shape and molecular weight. Thus, compound **16c**, an adenine analog of **10c** was synthesized according to the procedure depicted in Scheme 2.5, and was utilized as a standard (the synthesis of nucleoside **15** is outlined in the experimental section).

**Scheme 2.5.** The Synthesis of Compound **16c**.



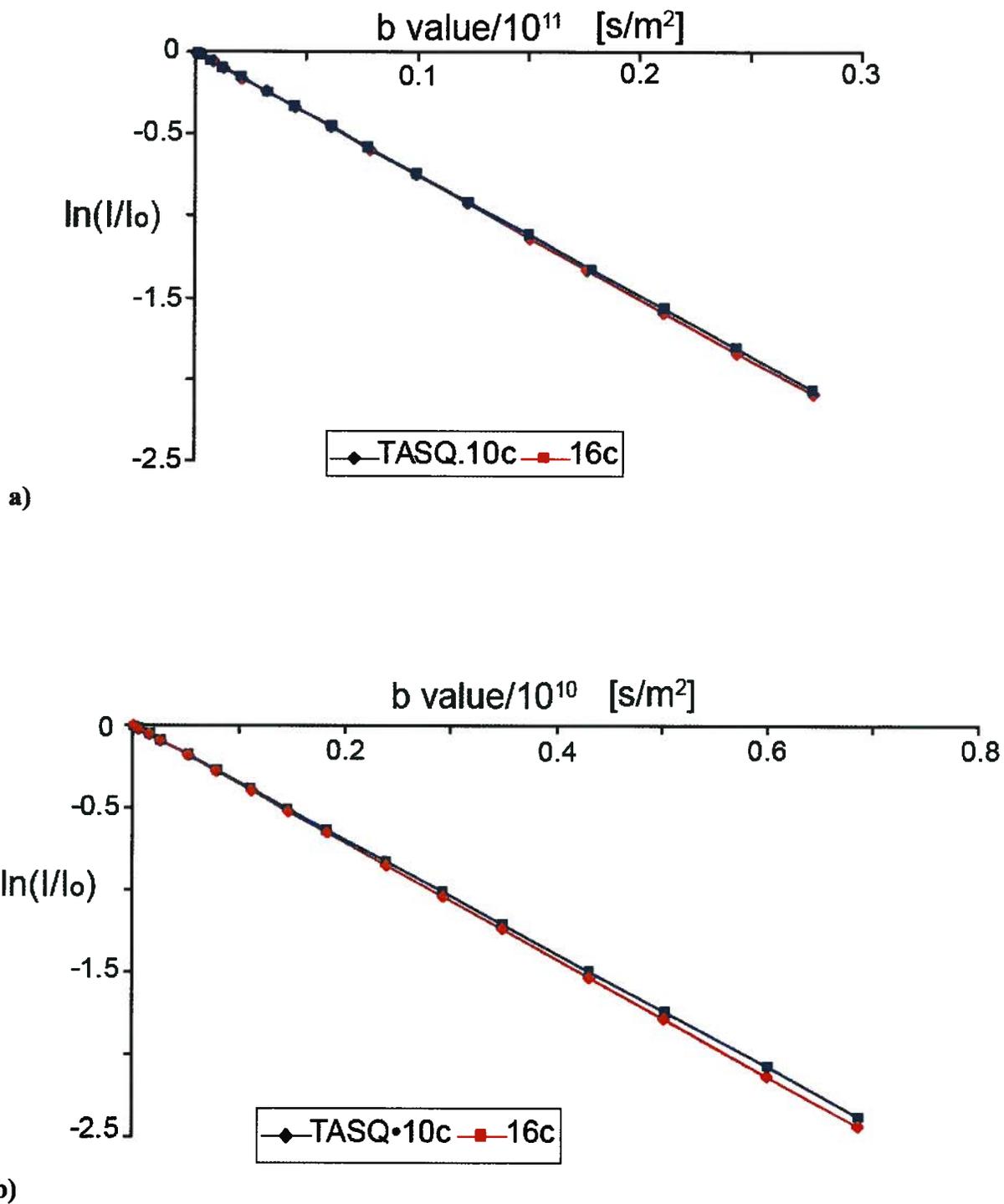
Mixtures of **10c-16c** in  $\text{CDCl}_3$  and in  $\text{DMSO-d}_6$  were prepared and subjected to analysis (Table 2.3). Both mixtures were in equimolar ratios and kept in the same NMR tube under identical conditions of concentration and temperature. In both cases, no interactions were observed between the standard and the analyzed sample. The experiments were performed at 295 K using a BPLED gradient pulse sequence, and the diffusion coefficients (D) were calculated by plotting the natural logarithm of the signal intensity versus the b-value for the  $\text{H1}'$  signal (Figure 2.17). The results are summarized in Table 2.3 and Figure 2.17. Figure 2.18 shows the exponential decay of the signals which are used for the calculation of diffusion coefficients as a function of the gradient strength (g).

**Table 2.3.** Diffusion coefficients (D) of **10c** and **16c** in  $\text{CDCl}_3$  and  $\text{DMSO-d}_6$  at 295 K.

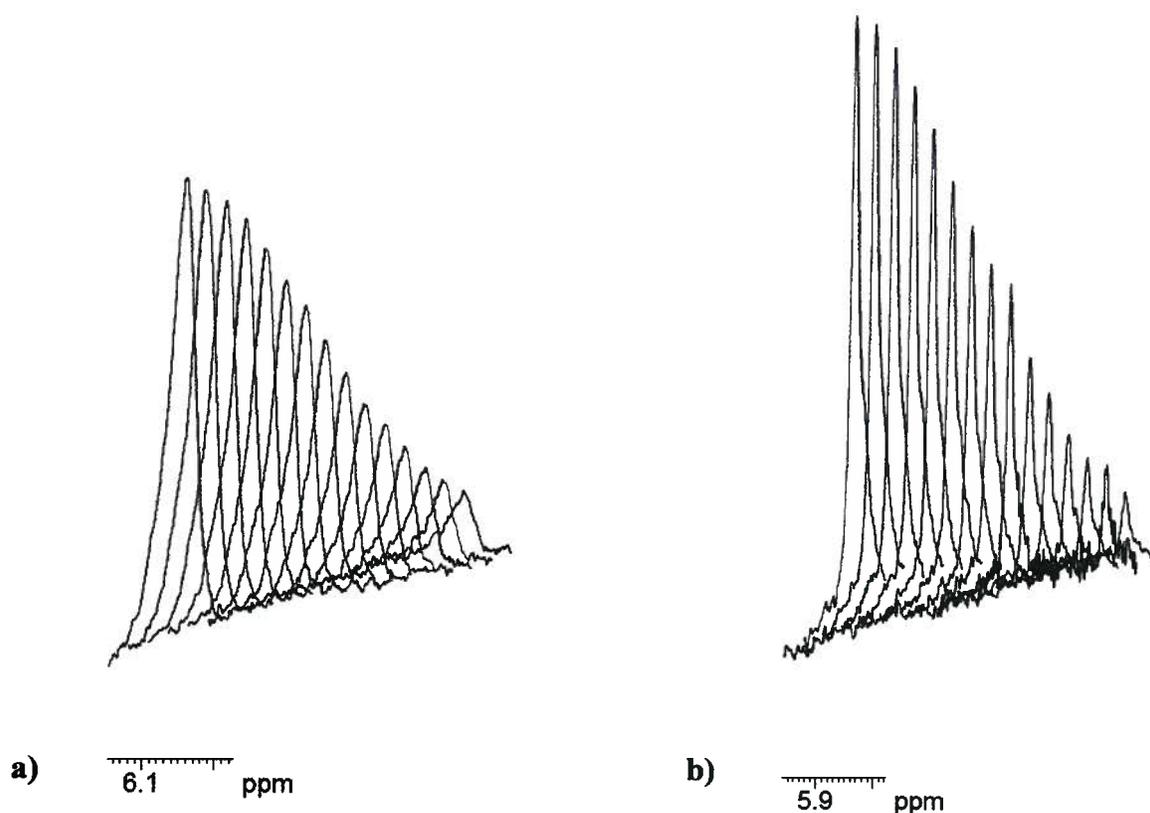
Solvent	$D_{16c}$ [ $\times 10^{-10} \text{ m}^2/\text{s}$ ]	$D_{10c}$ [ $\times 10^{-10} \text{ m}^2/\text{s}$ ]	% Ratio
$\text{DMSO-d}_6$	$0.75 \pm 0.02$	$0.74 \pm 0.02$	99
$\text{CDCl}_3$	$3.55 \pm 0.10$	$3.45 \pm 0.10$	97

From Figure 2.17a-b, it can be clearly seen that **10c** remains monomeric in both  $\text{DMSO-d}_6$  ( $\epsilon_r = 47$ ) and  $\text{CDCl}_3$  ( $\epsilon_r = 4.8$ ). The observed stoichiometry in  $\text{CDCl}_3$  is consistent with the previously determined basket-like geometry of TASQs, and rules out the presence of higher order aggregates.

**Figure 2.17.** Normalized signal decay as a function of the b value at 295 K for a) **10c-16c** in DMSO-d<sub>6</sub>, b) **10c-16c** in CDCl<sub>3</sub>. The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value =  $(2 \cdot \pi \cdot \gamma \cdot g \cdot \delta)^2 (\Delta - \delta/3) s/m^2$ ]



**Figure 2.18.** Normalized signal decay as a function of the  $g$  (gradient strength) at 295 K for a) the H1' signal of **10c** in DMSO- $d_6$ , and b) the H1' signal of **10c** in  $CDCl_3$ .



### 2.3.12 The Role of Water

So far we have provided a detailed analysis of the structure and stoichiometry of TASQs, and have highlighted the important role of cavitand in stabilization and formation of a cation-free TASQ. This section aims to evaluate whether the presence or absence of a trace amount of water in  $CDCl_3$  influences the formation and stabilization of TASQs or not. It has been proposed that water can take the place of cations in the centre of G-quartet, and stabilize the resulting structure.<sup>54</sup>

The results of variable-temperature  $^1\text{H}$  NMR experiments indicate that absorbed water in  $\text{CDCl}_3$  interacts with TASQ **10c** (Figure 2.10). The signal of H-bonded water molecules diminishes, and shows a downfield shift at lower temperatures. Presumably H-bonded water molecules experience a slow exchange and are more tightly bound at this temperature. Davis and colleagues have reported that water can stabilize a calix[4]arene-guanosine dimer.<sup>54</sup> While they have shown the relative location of water molecules in a sodium-bound dimer, the role and location of water in a cation-free system is still in question. We found that the presence of water is not crucial for the formation of soluble TASQs. **10c** is freely soluble in dry  $\text{CDCl}_3$ , and forms a G-quartet in anhydrous conditions. The changes in the chemical shift of exchangeable protons of this system in the presence or absence of water are very similar in variable-temperature  $^1\text{H}$  NMR experiments. Addition of water to a solution of TASQ **10c** in  $\text{CDCl}_3$  resulted in immediate precipitation. Also, no NOE was detected between water and guanosine protons at  $-20$  and  $-40$  °C.

It should be noted that the anhydrous condition in this experiment is defined as a condition under which no water signal is present in the  $^1\text{H}$  NMR spectra. In order to obtain this condition, both the sample and the NMR solvent have to be dry. The sample was dried *in vacuo* for 48 hours. The NMR solvent was dried following the literature procedure.<sup>55</sup>  $\text{CDCl}_3$  was stored over anhydrous potassium carbonate which is used to partially dry and deacidify the solvent. The solution obtained by this method contains 2-3 mM of residual water. The remaining water in the solution was then removed by the addition of a small amount of crushed 4 Å molecular sieves into the NMR tube. The prolonged storage (more than 48 hours) of the cation-free TASQ over molecular sieves is problematic as it results in cation-extraction and appearance of new signals.

## 2.4 Summary and Conclusions

In this Chapter, we have introduced the concept of template-assembled synthetic G-quartets (TASQs) as an approach for controlling the morphology of G-quadruplex assemblies. The ability to combine versatile structures, such as G-quadruplexes with rigid and shape-persistent templates such as cavitands may provide further scope for developments in biology, material science, and nanotechnology.

Following this strategy, the first generation of lipophilic TASQs were designed and synthesized by “click” chemistry. The incorporation of the cavitand template into the G-quartet system resulted in hybrid compounds which displayed discrete spectra in DMSO- $d_6$  and  $CDCl_3$ . The identity of the structure formed in  $CDCl_3$  was studied in detail using NMR spectroscopic methods, and was found to be a cation-free G-quartet. Variable-temperature  $^1H$  NMR experiments demonstrate that this G-quartet basket has considerable thermodynamic stability. In addition, the size of TASQ was evaluated by a series of PFG NMR experiments. These experiments characterized the TASQ as a unimolecular structure, and confirmed the basket-like morphology of it in solution. TASQs are the first unimolecular lipophilic G-quartets. They are also the first unimolecular G-quartets that do not require a cation for their structure (the proof of a cation-free system will be provided in Chapter 3 by both  $^1H$  NMR and Na-analysis).

Overall, these results illustrate the power of preorganization, and underscore the role of the cavitand in formation and stabilization of TASQs. The next Chapter will deal with the interaction of TASQs with cations.

## 2.5 Experimental

### 2.5.1 General

All chemicals were purchased from Sigma-Aldrich Chemical Co. Inc., and were used without further purification. *N,N*-dimethylformamide (DMF) was dried over 4 Å molecular sieves under a nitrogen atmosphere. Reagent grade THF was distilled under nitrogen from sodium benzophenone ketyl. Column chromatography was used for purification using silica gel (230-400 mesh, BDH) and silical gel glass backed analytical plates (0.2 mm) were used for thin layer chromatography (TLC) with UV detection. Deuterated solvents were purchased from Cambridge Isotope Laboratories. <sup>1</sup>H NMR spectra were recorded on Bruker AV-400inv, and AV-400dir spectrometers using the residual solvent signal as the reference. At temperatures other than ambient temperature, the NMR samples were equilibrated in the spectrometer for 20 minutes prior to data acquisition. 2D data were acquired for a 2×10<sup>-2</sup> M solution of samples. Phase sensitive 2D NOESY experiments were performed on an AV-400inv spectrometer. 256 time increments were collected (34 scans per increment) with mixing time of 100 ms. Three different mixing times were tested to check the spin diffusion limit. CD spectra were acquired using a JASCO J-710 or J-810. Each spectrum was the average of three scans corrected for the baseline. All spectra were obtained using a 1 mm path length cuvette. MALDI spectra were collected on a Bruker Biflex IV in reflectron mode using trans-3-indoleacrylic acid as the matrix. IR spectra were recorded on a Mattson Genesis spectrometer as thin film.

## 2.5.2 Synthesis of Cavitands 6a-c

Cavitands **6a-c** were synthesized according to the procedure described by Cram.<sup>19</sup> Cavitand **6c** has been reported by another group which obtained a similar yield.<sup>18</sup> Cavitand **6b** has been reported by Cram, but has not been fully characterized.<sup>19</sup> Cavitand **6a** has not been previously reported. The general procedure is as follows: Tetrol **5** (3 mmol; Scheme 2.2), potassium carbonate (6g, 44 mmol) and 5 mL 80% propargyl bromide in toluene (44 mmol) were added to 300 ml of acetone and were heated to reflux for 20 h. The solvent was evaporated and the residue was taken up in 300 ml of CH<sub>2</sub>Cl<sub>2</sub> and stirred. The resulting suspension was filtered and the filtrate was dried under reduced pressure and subjected to column chromatography on silica gel. The pure product was eluted with CH<sub>2</sub>Cl<sub>2</sub>-20% EtOAc/CH<sub>2</sub>Cl<sub>2</sub>.

**Cavitand 6c:** yield 68%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 7.28 (s, 4H, ArH), 5.87 (d, 4H, H<sub>out</sub>), 4.61 (d, 8H, O-CH<sub>2</sub>), 4.53 (q, 4H, H<sub>methine</sub>), 4.23 (d, 4H, H<sub>in</sub>), 3.49 (t, 4H, C≡CH), 2.30 (m, 8 H, CH<sub>2</sub> cavitand feet), 1.39 (m, 8 H, CH<sub>2</sub> feet) 1.23-1.30 (m, 64 H, CH<sub>2</sub> feet), 0.84 (t, 12 H, CH<sub>3</sub> feet) ppm.

**MS (ESI) *m/z*:** calculated: 1369.88 for C<sub>88</sub>H<sub>120</sub>O<sub>12</sub>; found: [M+H<sup>+</sup>] 1371.1.

**Cavitand 6b:** yield 71%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 7.36 (s, 4H, ArH), 7.20 (m, 20H, phenyl), 5.90 (d, 4H, H<sub>out</sub>), 4.66 (d, 8H, O-CH<sub>2</sub>), 4.61 (q, 4H, H<sub>methine</sub>), 4.25 (d, 4H, H<sub>in</sub>), 3.53 (t, 4H, C≡CH), 2.60 (m, 16 H, CH<sub>2</sub> cavitand feet) ppm.

**MS (ESI) *m/z*:** calculated: 1168.44 for C<sub>76</sub>H<sub>64</sub>O<sub>12</sub>; found: [M+H<sup>+</sup>] 1169.6.

**Cavitand 6a:** yield 69%,  $^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ , 298 K)  $\delta$  7.46 (s, 4H, ArH), 5.90 (d, 4H,  $\text{H}_{\text{out}}$ ), 4.78 (q, 4H,  $\text{H}_{\text{methine}}$ ), 4.65 (d, 8H, O- $\text{CH}_2$ ), 4.26 (d, 4H,  $\text{H}_{\text{in}}$ ), 3.51 (t, 4H,  $\text{C}\equiv\text{CH}$ ), 1.80 (d, 12H,  $\text{CH}_3$ ) ppm.

**MS (ESI)  $m/z$ :** calculated: 808.25 for  $\text{C}_{48}\text{H}_{40}\text{O}_{12}$ ; found:  $[\text{M}+\text{H}^+]$  809.1.

### 2.5.3 Synthesis of 5'-Azido-2',3'-*O*-Isopropylidene Guanosine 9

Compound **9** was obtained by a known method,<sup>20</sup> and the characterization data matched the literature values.

**Compound 9:** yield 42%,  $^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ , 298 K)  $\delta$  10.72 (s, 1H, NH), 7.90 (d, 1H, H8), 6.56 (bs, 2H,  $\text{NH}_2$ ), 6.01 (d, 1H,  $\text{H}1'$ ), 5.30 (dd, 1H,  $\text{H}2'$ ), 5.06 (dd, 1H,  $\text{H}3'$ ), 4.23 (m, 1H,  $\text{H}4'$ ), 3.58 (dd, 1H,  $\text{H}5'$ ), 3.54 (dd, 1H,  $\text{H}5'$ ), 1.52 (s, 3H,  $\text{CH}_3$ ), 1.32 (s, 3H,  $\text{CH}_3$ ) ppm.

**MS (ESI)  $m/z$ :** calculated: 348.13 for  $\text{C}_{13}\text{H}_{16}\text{N}_8\text{O}_4$ ; found:  $[\text{M}+\text{H}^+]$  349.1.

### 2.5.4 Synthesis of TASQs 10a-c

**General procedure.** To a solution of cavitand **6a-c** (0.081 mmol) and azide **9** (125 mg, 0.360 mmol) in 10 mL of degassed DMSO was added  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (0.202 mL, 0.040 M aqueous solution, 0.008 mmol) and sodium ascorbate (0.202 mL of freshly prepared 0.400 M

aqueous solution, 0.081 mmol). The reaction mixture was stirred at 60 °C overnight under a positive pressure of N<sub>2</sub>. The solvent was removed *in vacuo*, and the crude products were washed thoroughly with hot de-ionized water. A few drops of ammonium hydroxide were used to remove the copper catalyst. The crude reaction mixture was embedded on silica gel and the unreacted azide eluted with a 1:3 mixture of EtOH/EtOAc. The pure product was collected by washing the column with THF. The solvent was evaporated and the obtained precipitate dissolved in DMSO. Precipitation with de-ionized water, filtration, and subsequent washing of the precipitate with hot de-ionized water yielded the pure product.

**Compound 10c:** yield 67%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 10.78 (s, 4 H, NH), 8.10 (s, 4 H, H<sub>c</sub>), 7.82 (s, 4 H, H<sub>8</sub>), 7.26 (s, 4 H, ArH), 6.59 (bs, 8H, NH<sub>2</sub>), 6.03 (d, 4H, H1'), 5.88 (d, 4 H, H<sub>out</sub>), 5.27 (dd, 4 H, H2'), 5.24 (dd, 4 H, H3'), 4.90 (m, 8 H, H<sub>a</sub>, H<sub>b</sub>), 4.69 (m, 8 H, H5'a, H5'b), 4.55 (t, 4 H, H<sub>d</sub>), 4.48 (m, 4 H, H4'), 4.24 (d, 4 H, H<sub>in</sub>), 2.30 (m, 8 H, CH<sub>2</sub> cavitand feet), 1.48 (s, 12 H, CH<sub>3</sub> sugar), 1.39 (m, 8 H, CH<sub>2</sub> feet), 1.29 (s, 12 H, CH<sub>3</sub> sugar), 1.23-1.30 (m, 64 H, CH<sub>2</sub> feet), 0.84 (t, 12 H, CH<sub>3</sub> feet). <sup>13</sup>C NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 156.8, 154.2, 153.7, 150.3, 147.5, 143.8, 143.4, 138.7, 137.9, 136.5, 124.7, 117.1, 115.9, 113.4, 88.6, 85.1, 83.8, 81.7, 66.5, 51.4, 36.9, 31.3, 29.3, 29.2, 29.1, 28.8, 27.7, 26.8, 25.2, 25.2, 22.1, 13.8.

**MS (MALDI-TOF) *m/z*:** calculated: 2785.40 for C<sub>140</sub>H<sub>184</sub>N<sub>32</sub>O<sub>28</sub>·Na; found: [M+Na<sup>+</sup>] 2785.2.

**Compound 10b:** yield 62%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 10.79 (s, 4 H, NH), 8.11 (s, 4 H, H<sub>c</sub>), 7.81 (s, 4 H, H<sub>8</sub>), 7.33 (s, 4 H, ArH), 7.15-7.26 (m, 20 H, Ar), 6.59 (bs, 8H, NH<sub>2</sub>), 6.02 (d, 4H, H1'), 5.89 (d, 4 H, H<sub>out</sub>), 5.27 (dd, 4 H, H2'), 5.24 (dd, 4 H, H3'), 4.93 (m, 8 H, H<sub>a</sub>, H<sub>b</sub>), 4.70 (m, 8 H, H5'a, H5'b), 4.63 (t, 4 H, H<sub>d</sub>), 4.49 (m, 4 H, H4'), 4.27 (d, 4 H, H<sub>in</sub>), 2.58 (m,

16 H, CH<sub>2</sub> cavitand feet), 1.47 (s, 12 H, CH<sub>3</sub> sugar), 1.29 (s, 12 H, CH<sub>3</sub> sugar). <sup>13</sup>C NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 156.8, 153.7, 150.3, 147.6, 143.9, 143.4, 141.6, 138.7, 136.4, 128.7, 128.2, 125.8, 124.8, 117.1, 116.0, 113.4, 99.3, 88.6, 85.0, 83.7, 81.6, 66.5, 51.4, 36.9, 33.8, 31.3, 26.8, 25.2.

**MS (MALDI-TOF) *m/z***: calculated: 2584.96 for C<sub>128</sub>H<sub>128</sub>N<sub>32</sub>O<sub>28</sub>·Na; found: [M+Na<sup>+</sup>] 2584.3.

**Compound 10a**: yield 65%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 10.79 (s, 4 H, NH), 8.10 (s, 4 H, H<sub>c</sub>), 7.83 (s, 4 H, H<sub>8</sub>), 7.42 (s, 4 H, ArH), 6.59 (bs, 8H, NH<sub>2</sub>), 6.04 (d, 4H, H1'), 5.88 (d, 4 H, H<sub>out</sub>), 5.28 (dd, 4 H, H2'), 5.24 (dd, 4 H, H3'), 4.91 (m, 8 H, H<sub>a</sub>, H<sub>b</sub>), 4.75 (q, 4 H, H<sub>d</sub>), 4.70 (m, 8 H, H5'a, H5'b), 4.48 (m, 4 H, H4'), 4.25 (d, 4 H, H<sub>in</sub>), 1.78 (d, 12 H, CH<sub>3</sub> cavitand feet), 1.48 (s, 12 H, CH<sub>3</sub> sugar), 1.30 (s, 12 H, CH<sub>3</sub> sugar). <sup>13</sup>C NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 156.8, 153.7, 150.3, 147.1, 143.5, 143.3, 139.7, 136.5, 124.8, 117.1, 115.6, 113.4, 99.2, 88.6, 85.1, 83.7, 81.6, 66.4, 66.4, 51.4, 31.2, 26.8, 25.2, 15.7.

**MS (MALDI-TOF) *m/z***: calculated: 2224.77 for C<sub>100</sub>H<sub>104</sub>N<sub>32</sub>O<sub>28</sub>·Na; found: [M+Na<sup>+</sup>] 2224.5.

### 2.5.5 Synthesis of 5'-Azido-2',3'-*O*-Isopropylidene Adenosine 15

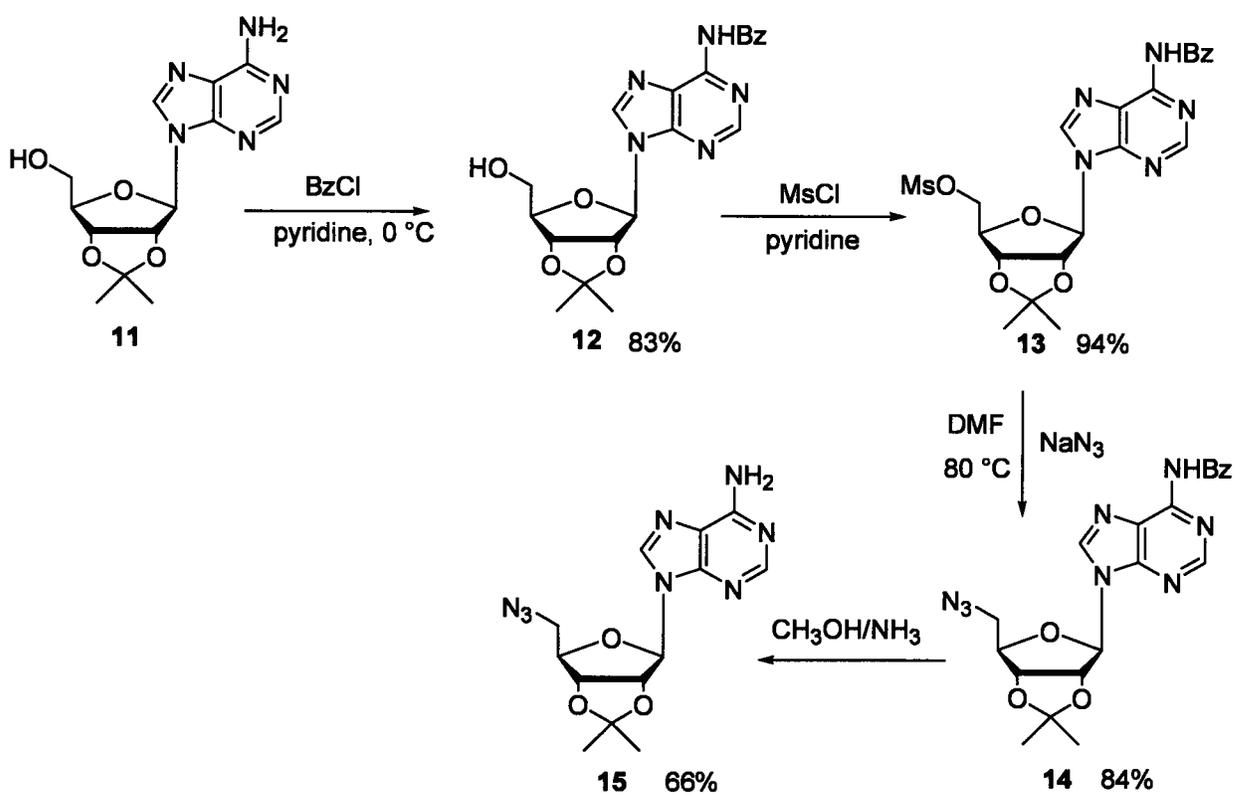
Synthesis of compound **15** was accomplished using a method similar to the synthesis of compound **9** (Scheme 2.6).<sup>56</sup> Briefly, the exocyclic amino group of **11** was protected, and the product was reacted with methanesulfonyl chloride to give compound **13**. This product was then

treated with sodium azide, and deprotected to yield the final product. The  $^1\text{H}$  NMR data matched the literature values.

**Compound 15:** yield 66%,  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.40 (s, 1H, H2), 7.95 (s, 1H, H8), 6.12 (d, 1H, H1'), 5.62 (bs, 2H,  $\text{NH}_2$ ), 5.45 (dd, 1H, H2'), 5.05 (dd, 1H, H3'), 4.39 (m, 1H, H4'), 3.61 (dd, 1H, H5'), 3.56 (dd, 1H, H5'), 1.61 (s, 3H,  $\text{CH}_3$ ), 1.42 (s, 3H,  $\text{CH}_3$ ) ppm.

**MS (ESI)  $m/z$ :** calculated: 332.13 for  $\text{C}_{13}\text{H}_{16}\text{N}_8\text{O}_3$ ; found:  $[\text{M}+\text{H}^+]$  333.3.

**Scheme 2.6.** The Synthesis of Nucleoside 15.



## 2.5.6 Synthesis of Compound 16c

The adenine analog of TASQ **10c** was synthesized following the general procedure described in the section 2.5.4, and was utilized for diffusion NMR and cation-selectivity studies. To a solution of cavitand **6c** (111 mg, 0.081 mmol) and azide **15** (120 mg, 0.360 mmol) in 10 mL of degassed DMSO was added CuSO<sub>4</sub>·5H<sub>2</sub>O (0.202 mL, 0.040 M aqueous solution, 0.008 mmol) and sodium ascorbate (0.202 mL of freshly prepared 0.400 M aqueous solution, 0.081 mmol). The reaction mixture was stirred at 60 °C overnight under a positive pressure of N<sub>2</sub>. The solvent was removed *in vacuo*, and the crude products were washed thoroughly with hot de-ionized water. A few drops of ammonium hydroxide were used to remove the copper catalyst. The crude reaction mixture was subjected to column chromatography on silica gel using a 1:3 mixture of EtOH/EtOAc. The solvent was evaporated and the obtained precipitate dissolved in DMSO. Precipitation with de-ionized water, filtration, and subsequent washing of the precipitate with hot de-ionized water yielded the pure product.

**Compound 16c:** yield 68%, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 8.26 (s, 4 H, H<sub>8</sub>), 8.18 (s, 4 H, H<sub>e</sub>), 8.06 (s, 4 H, H<sub>c</sub>), 7.33 (bs, 8H, NH<sub>2</sub>), 7.26 (s, 4 H, ArH), 6.22 (d, 4H, H1'), 5.88 (d, 4 H, H<sub>out</sub>), 5.46 (dd, 4 H, H2'), 5.14 (dd, 4 H, H3'), 4.90 (m, 8 H, H<sub>a</sub>, H<sub>b</sub>), 4.76 (dd, 4 H, H5'b), 4.65 (dd, 4 H, H5'a), 4.56 (t, 4 H, H<sub>d</sub>), 4.55 (m, 4 H, H4'), 4.25 (d, 4 H, H<sub>in</sub>), 2.30 (m, 8 H, CH<sub>2</sub> cavitand feet), 1.49 (s, 12 H, CH<sub>3</sub> sugar), 1.38 (m, 8 H, CH<sub>2</sub> feet), 1.29 (s, 12 H, CH<sub>3</sub> sugar), 1.20-1.28 (m, 64 H, CH<sub>2</sub> feet), 0.84 (t, 12 H, CH<sub>3</sub> feet). <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 156.2, 153.2, 148.9, 148.2, 148.1, 145.1, 144.0, 140.0, 139.0, 124.3, 120.0, 114.9, 114.7, 99.6, 90.4, 84.9, 83.7, 81.7, 67.4, 51.3, 37.0, 32.0, 30.0, 29.8 (multiple peaks), 29.5, 28.1, 27.2, 25.5, 22.8, 14.2.

**MS (MALDI-TOF)  $m/z$ :** calculated: 2699.42 for  $C_{140}H_{184}N_{32}O_{24}.H^+$ ; found:  $[M+H^+]$  2700.1.

## **2.5.7 PFG NMR Spectroscopic Experiments**

### **2.5.7.1 General**

All diffusion NMR experiments were carried out on a Bruker AV-400inv spectrometer equipped with a 5mm BBI Z-gradient probe (inverse broadband probe with z-gradient coil). All measurements were performed using a BPLED gradient pulse sequence (ledbpgp2s). The length of the diffusion gradient was optimized for each diffusion time to obtain at least 95% signal attenuation due to diffusion. Eddy current ( $t_e$ ) was set at 5 ms. All spectra were acquired by using Shigemi tubes (Shigemi, Inc., Allison Park, PA) to minimize the effect of convection. All measurements were performed at 22 °C on a  $3.6 \times 10^{-3}$  M solution of samples, unless otherwise noted. A total of 160  $\mu$ L of solution was used for each analysis. The spectra were processed and plotted by TopSpin 1.3 and TopSpin Plot Editor 1.3 respectively. All data are means of at least three measurements.

### **2.5.7.2 Experimental (Bruker) Parameters for PFG NMR of 10c-16c in $CDCl_3$**

NS [scan]: 32	TD (F2): 8K	TD (F1): 16
D21 [delay]: 5ms	D20 [ $\Delta$ ]: 100 ms	DS: 8
P30 [ $\delta/2$ ]: 1.5 ms	P19: 0.75 ms	GPZ6 [G]: 100%
GPZ7: -17.13%	GPZ8: -13.17	

Gradients were calibrated from 2% up to 95% in 16 linear steps (TD1).

### **2.5.7.3 Experimental (Bruker) Parameters for PFG NMR of 10c-16c in DMSO-d<sub>6</sub>**

NS [scan]: 16	TD (F2): 8K	TD (F1): 16
D21 [delay]: 5ms	D20 [ $\Delta$ ]: 150 ms	DS: 8
P30 [ $\delta/2$ ]: 2.5 ms	P19: 1.25 ms	GPZ6 [G]: 100%
GPZ7: -17.13%	GPZ8: -13.17	

Gradients were calibrated from 2% up to 95% in 16 linear steps (TD1).

## 2.6 References

- (1) Keniry, M. A. *Biopolymers* **2000**, *56*, 123-146.
- (2) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.
- (3) Giorgi, T.; Grepioni, F.; Manet, I.; Mariani, P.; Masiero, S.; Mezzina, E.; Pieraccini, S.; Saturni, L.; Spada, G. P.; Gottarelli, G. *Chem.-Eur. J.* **2002**, *8*, 2143-2152.
- (4) Li, Z.; Mirkin, C. A. *J. Am. Chem. Soc.* **2005**, *127*, 11568-11569.
- (5) Calzolari, A.; Di Felice, R.; Molinari, E.; Garbesi, A. *Appl. Phys. Lett.* **2002**, *80*, 3331-3333.
- (6) Vo, T. U.; McGown, L. B. *Electrophoresis* **2004**, *25*, 1230-1236.
- (7) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004-2021.
- (8) Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36*, 1249-1262.
- (9) Padwa, A. *1, 3-dipolar cycloaddition chemistry*; Wiley: New York; Chichester, 1984.
- (10) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596-2599.
- (11) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.
- (12) Creutz, C. *Inorg. Chem.* **1981**, *20*, 4449-4452.
- (13) Dondoni, A.; Marra, A. *J. Org. Chem.* **2006**, *71*, 7546-7557.
- (14) Ryu, E. H.; Zhao, Y. *Org. Lett.* **2005**, *7*, 1035-1037.
- (15) Kaval, N.; Ermolat'ev, D.; Appukkuttan, P.; Dehaen, W.; Kappe, C. O.; Van der Eycken, E. *J. Comb. Chem.* **2005**, *7*, 490-502.
- (16) Malkoch, M.; Schleicher, K.; Drockenmüller, E.; Hawker, C. J.; Russell, T. P.; Wu, P.; Fokin, V. V. *Macromolecules* **2005**, *38*, 3663-3678.
- (17) Moorhouse, A. D.; Santos, A. M.; Gunaratnam, M.; Moore, M.; Neidle, S.; Moses, J. E. *J. Am. Chem. Soc.* **2006**, *128*, 15972-15973.
- (18) Barrett, E. S.; Irwin, J. L.; Picker, K.; Sherburn, M. S. *Aust. J. Chem.* **2002**, *55*, 319-325.
- (19) Cram, D. J.; Jaeger, R.; Deshayes, K. *J. Am. Chem. Soc.* **1993**, *115*, 10111-10116.

- (20) Stout, M. G.; Robins, M. J.; Olsen, R. K.; Robins, R. K. *J. Med. Chem.* **1969**, *12*, 658-662.
- (21) Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010-5017.
- (22) Gottarelli, G.; Masiero, S.; Mezzina, E.; Spada, G. P.; Mariani, P.; Recanatini, M. *Helv. Chim. Acta* **1998**, *81*, 2078-2092.
- (23) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164-168.
- (24) Marlow, A. L.; Mezzina, E.; Spada, G. P.; Masiero, S.; Davis, J. T.; Gottarelli, G. *J. Org. Chem.* **1999**, *64*, 5116-5123.
- (25) Liu, X. Y.; Kwan, I. C. M.; Wang, S. N.; Wu, G. *Org. Lett.* **2006**, *8*, 3685-3688.
- (26) Wang, Y.; Patel, D. J. *Biochemistry* **1992**, *31*, 8112-8119.
- (27) Feigon, J.; Wang, A. H. J.; Vandermarel, G. A.; Vanboom, J. H.; Rich, A. *Nucleic Acids Res.* **1984**, *12*, 1243-1263.
- (28) Patel, D. J.; Kozlowski, S. A.; Nordheim, A.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 1413-1417.
- (29) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A. *Angew. Chem., Int. Ed.* **2000**, *39*, 1300-1303.
- (30) Aboulela, F.; Murchie, A. I. H.; Lilley, D. M. *Nature* **1992**, *360*, 280-282.
- (31) Functionalization of the cavitand with other purine bases such as adenine did not lead to similar changes in the chemical shift of non-exchangable protons. The H<sub>out</sub> proton of **16c** (the adenine analog of **10c**, Scheme 2.5) resonates at 5.7 ppm in CDCl<sub>3</sub>, a value pretty close to the chemical shift of this proton in DMSO-d<sub>6</sub>.
- (32) Williams, L. D.; Williams, N. G.; Shaw, B. R. *J. Am. Chem. Soc.* **1990**, *112*, 829-833.
- (33) Kyogoku, Y.; Lord, R. C.; Rich, A. *Science* **1966**, *154*, 518-520.
- (34) Fasman, G. D. *Circular dichroism and the conformational analysis of biomolecules*; Plenum Press: New York, 1996.
- (35) Forman, S. L.; Fettingner, J. C.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067.
- (36) Gottarelli, G.; Masiero, S.; Spada, G. P. *Enantiomer* **1998**, *3*, 429-438.
- (37) Crank, J. *The mathematics of diffusion*; 2nd ed. ed.; Clarendon Press: Oxford, 1979.
- (38) Einstein, A. *Ann. Phys.* **1906**, *19*, 298-306.

- (39) Cussler, E. L. *Diffusion: mass transfer in fluid systems*; 2nd ed. ed.; Cambridge University Press: Cambridge, 1997.
- (40) Otting, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *31*, 259-285.
- (41) Grant, D. M.; Harris, R. K. *Encyclopedia of nuclear magnetic resonance*; Wiley: New York; Chichester, 2002.
- (42) Waldeck, A. R.; Kuchel, P. W.; Lennon, A. J.; Chapman, B. E. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *30*, 39-68.
- (43) Cohen, Y.; Avram, L.; Frish, L. *Angew. Chem., Int. Ed.* **2005**, *44*, 520-554.
- (44) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288-&.
- (45) Kazanis, S. *Measuring Diffusion by NMR*; Bruker Instruments Inc., 2000.
- (46) Price, W. S. *Concepts Magn. Reson.* **1997**, *9*, 299-336.
- (47) Price, W. S. *Concepts Magn. Reson.* **1998**, *10*, 197-237.
- (48) Schalley, C. A. *Analytical methods in supramolecular chemistry*; Weinheim: Wiley-VCH, 2007.
- (49) Evan-Salem, T.; Baruch, I.; Avram, L.; Cohen, Y.; Palmer, L. C.; Rebek, J. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12296-12300.
- (50) Avram, L.; Cohen, Y. *J. Am. Chem. Soc.* **2002**, *124*, 15148-15149.
- (51) Kaucher, M. S.; Lam, Y. F.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *Chem.-Eur. J.* **2004**, *11*, 164-173.
- (52) Evan-Salem, T.; Frish, L.; van Leeuwen, F. W. B.; Reinhoudt, D. N.; Verboom, W.; Kaucher, M. S.; Davis, J. T.; Cohen, Y. *Chem.-Eur. J.* **2007**, *13*, 1969-1977.
- (53) Timmerman, P.; Weidmann, J. L.; Jolliffe, K. A.; Prins, L. J.; Reinhoudt, D. N.; Shinkai, S.; Frish, L.; Cohen, Y. *J. Chem. Soc., Perkin Trans. 2* **2000**, 2077-2089.
- (54) Kotch, F. W.; Sidorov, V.; Lam, Y. F.; Kayser, K. J.; Li, H.; Kaucher, M. S.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 15140-15150.
- (55) Bonarlaw, R. P.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1995**, *117*, 259-271.
- (56) Ceulemans, G.; Vandendriessche, F.; Rozenski, J.; Herdewijn, P. *Nucleosides Nucleotides* **1995**, *14*, 117-127.

## **CHAPTER 3: Cation-Complexation Behavior of TASQs: The Preparation and Structural Characterization of TASQ·Cation Assemblies\***

### **3.1 Introduction**

Chapter 2 introduced the design and synthesis of TASQs. It illustrated the important role of the cavitand template in the assembly of a unimolecular G-quartet. This Chapter will describe the preparation and characterization of a series of TASQ·cation assemblies, and illustrate the effect of cations on the structure of TASQs.

### **3.2 Rationale for the Study of TASQ·Cation Assemblies**

Cations facilitate the formation of G-quartets and contribute to the stability and polymorphism of G-quadruplex structures.<sup>1-5</sup> They can induce structural changes or trigger conformational transitions. Therefore, numerous studies have been conducted on various aspects of cation-binding such as selectivity,<sup>6-8</sup> location,<sup>9-11</sup> exchange<sup>12-14</sup> and cation-dependent polymorphism.<sup>15,16</sup>

Unlike lipophilic and nucleic acid G-quadruplexes which are templated by cations, a TASQ is a singular G-quartet which is stabilized by a cavitand template. Owing to its particular shape, TASQs contain a binding pocket which is preorganized for cation binding. This places it

---

\* "A version of this Chapter has been submitted for publication. Nikan, M.; Sherman, J. C., Cation-Complexation Behavior of TASQs: The Emergence of Isolated G-Quartets."

in the same category as synthetic ionophores (e.g. spherands and crown ethers) and gives it the potential ability to interact with cations. In addition, TASQs contain aromatic surfaces which can participate in  $\pi$ -stacking. The cooperation of metal coordination, stacking and H-bonding interactions is expected to influence the morphology of the TASQs; however, the number of topologies that can be adopted by a TASQ will be limited by the presence of an external template. The major goal of this study is to determine the effect of cations on the structure of TASQs. Some of the questions that we are seeking to answer are: Is a TASQ capable of interacting with cations? If yes, does the TASQ maintain the G-quartet structure after coordination to cations? What are the sizes of the complexes formed? Are the TASQs selective for a particular cation? The answers to these questions will allow us to have a better understanding of the behavior of TASQs and to tailor the TASQs for a particular application in the future.

The natural and synthetic DNA interacting agents containing a binding site for cations are gaining lots of attention.<sup>17</sup> The concentration of ions in cancerous cells and in the vicinity of DNA is generally high.<sup>18</sup> Thus, it has been proposed that a combination of DNA-binding and cation-binding domains might help target the cancerous cells selectively.<sup>17</sup> In this regard TASQ might hold promise in therapeutic applications.

Our present work also provides more insight into the formation and structural modification of isolated G-quartets. Isolated G-quartets are rare and consequently little is known concerning their behavior. Attempts to capture an isolated G-quartet by linking the guanine bases inside a quartet have not been successful.<sup>19</sup> It has been reported that a single G-quartet is not an efficient cation-extractor.<sup>8</sup> TASQs bring new perspective and prove that efficient ionophoric systems can be prepared using the properties of only one G-quartet.

### 3.3 Results and Discussion

Four different cations, Na<sup>+</sup> (1.02 Å), K<sup>+</sup> (1.38 Å), Sr<sup>2+</sup> (1.18 Å) and Cs<sup>+</sup> (1.70 Å) were chosen for recognition analysis by TASQs based on their sizes, valencies, and G-quartet binding properties.<sup>20</sup> The first two cations are commonly found in G-quadruplex assemblies, and the Sr<sup>2+</sup> cation is known to bind effectively to G-quadruplexes.<sup>6</sup> With respect to G-quadruplexes, the cesium cation is generally considered to be a nonconstructing ion: the large size of the Cs<sup>+</sup> cation prevents proper stacking of the bases, and results in weakening of the G-quadruplex.<sup>21</sup> Thus, it usually serves as a capping ion.<sup>22</sup>

#### 3.3.1 Synthesis of TASQs

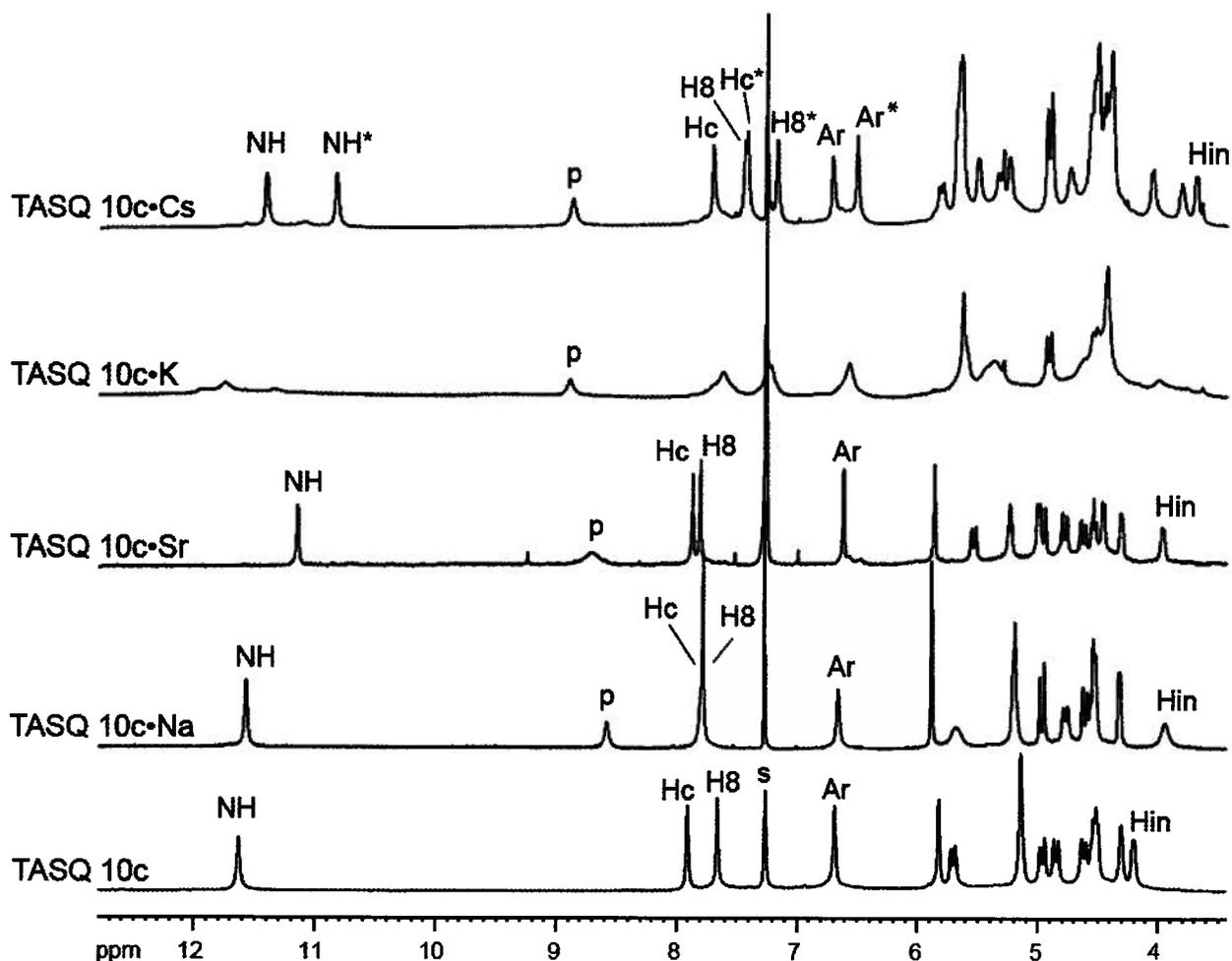
Our strategy for the synthesis of TASQs **10a-c** has been previously described in the section 2.3.4, where the key step is a “click” reaction between guanosine **9** and cavitands **6a-c**. The tetra-substituted products were obtained in good yield with no sign of the mono, di, or tri-substituted products.

#### 3.3.2 <sup>1</sup>H NMR Spectroscopic Studies

A series of solid-liquid extraction experiments were carried out on TASQ **10c** in CHCl<sub>3</sub> and the solid picrate salt of the sample cations. The products were subjected to <sup>1</sup>H NMR spectroscopic analysis in CDCl<sub>3</sub> (Figure 3.1). In all cases, the cation-complexation was

complete, resulting in the formation of a new set of signals and the complete loss of the signals due to the cation-free TASQ. A comparison of the spectra revealed that each assembly has its own unique spectrum that depends on the identity of the corresponding cation. As illustrated in Figure 3.1, small cations such as  $\text{Na}^+$  and  $\text{Sr}^{2+}$  give well-resolved spectra comprised of only one set of signals, while  $\text{K}^+$  yields a broad spectrum and  $\text{Cs}^+$  displays two distinct sets of signals.

**Figure 3.1.**  $^1\text{H}$  NMR spectra of TASQ-cation assemblies in  $\text{CDCl}_3$  at 400 MHz at 25 °C. The region between 6 and 12 ppm's has been assigned to make comparison easier. The signals labeled "p" and "s" indicate picrate and solvent respectively. TASQ 10c- $\text{Cs}^+$  gave two sets of signals, one of which has been marked with asterisks.



One common feature found in all of these spectra is that the H<sub>in</sub> proton of the cavitand is sensitive to the presence of cations. This sensitivity can be used as a probe for cation-binding and suggests that the coordinated cations are located near or inside the cavitand template (Figure 3.1). The following sections will discuss the characterization of each assembly in detail.

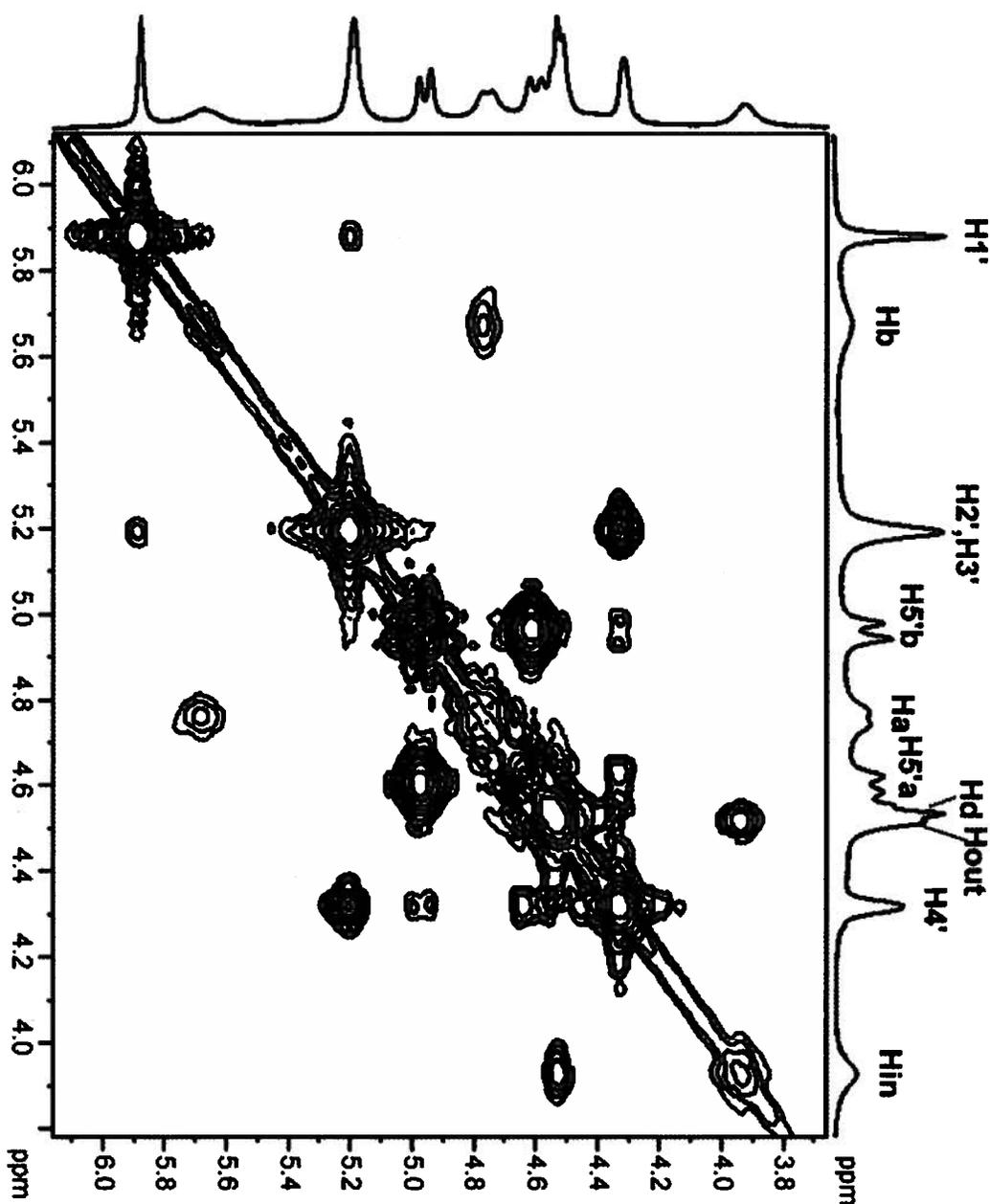
### 3.3.3 TASQ 10c·Na<sup>+</sup>

Due to its physiological importance, sodium-bound complexes are among the most studied G-quadruplexes in hydrophilic systems.<sup>1,23</sup> The sodium ion is a small cation which can occupy positions in the middle or between the quartets and move in or out of a G-quartet like it does in ion channels.<sup>2</sup> Unlike hydrophilic systems, sodium-bound G-quadruplexes in lipophilic systems have not been extensively studied. Most of the lipophilic studies have been performed on potassium-bound complexes.

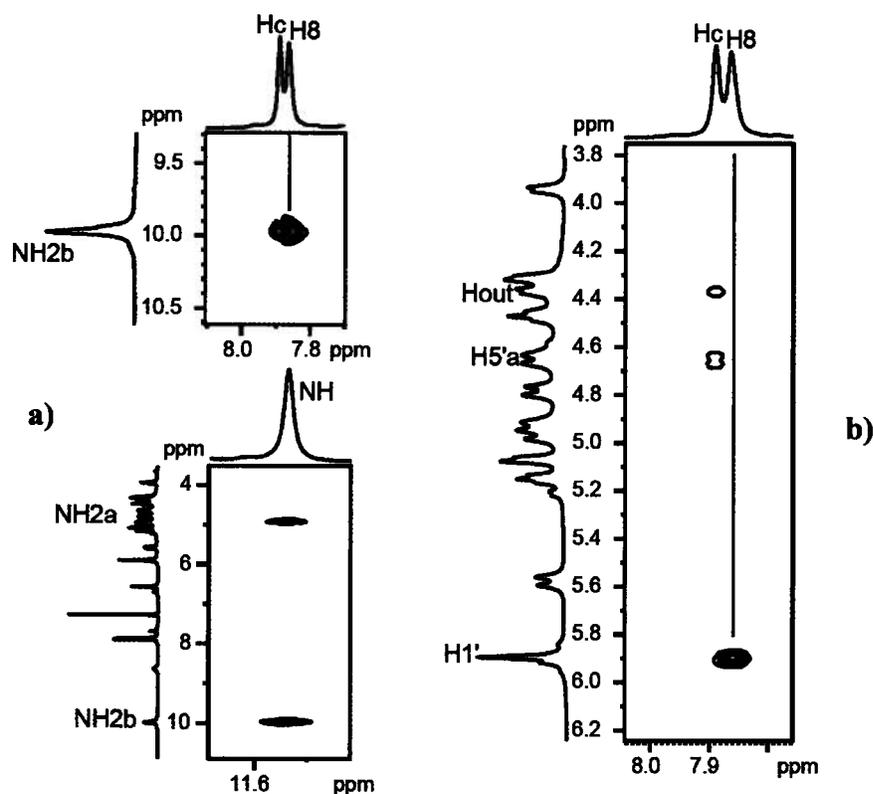
TASQ 10c binds sodium and yields <sup>1</sup>H NMR spectra characterized by only one set of signals. The following experiments were undertaken to analyze the morphology of the resulting complex. First, the signal assignments were derived from <sup>1</sup>H-<sup>1</sup>H COSY experiments (Figure 3.2). Then, a set of 2D NOESY experiments at -40 °C were acquired to investigate the H-bonding of the bases. The results of both experiments were consistent with the formation of a G-quartet in solution. TASQ 10c·Na<sup>+</sup> displayed all the common and specific features of the G-quartet formation that we have previously reported for the cation-free TASQs, including: 1) H-bonding of the amino groups which is confirmed by the downfield shift of the H-bonded amino signal (NH<sub>2b</sub>) and observed NOEs between NH/NH<sub>2b</sub> and NH/NH<sub>2a</sub>(Figure 3.3),<sup>24,25</sup> 2) NOE correlations between H8 and NH<sub>2b</sub><sup>26,27</sup> (Figure 3.3), and 3) changes in the chemical shifts of the

non-exchangeable protons with respect to the spectra in DMSO-d<sub>6</sub> (Figure 3.2).<sup>28</sup> 2D NOESY analysis revealed that the guanine bases in TASQ 10c·Na<sup>+</sup> have been driven into *syn* conformations which provide a suitable geometry for Hoogsteen pairing and forces the structure into a G-quartet (Figure 3.3b).<sup>29</sup>

**Figure 3.2.** <sup>1</sup>H-<sup>1</sup>H COSY of TASQ 10c·Na<sup>+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.

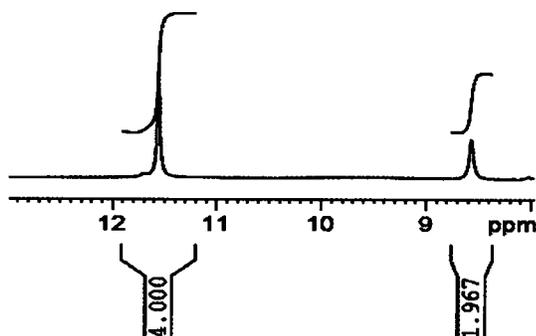


**Figure 3.3.** TASQ 10c·Na<sup>+</sup> shows NOEs similar to those shown by the cation-free system. These effects are indicative of a) the formation of G-quartet and, b) the syn conformation at 400 MHz in CDCl<sub>3</sub> at -40 °C. The NH<sub>2</sub> signals which appear broad at RT, sharpen up and become clearly visible at low temperature.



Although <sup>1</sup>H NMR, <sup>1</sup>H-<sup>1</sup>H COSY and <sup>1</sup>H-<sup>1</sup>H NOESY provided valuable insight into the structure of TASQ 10c·Na<sup>+</sup>, and confirm the presence of a G-quartet motif, they were not sufficient for the determination of the stoichiometry. The ratio of cation per quartet for this assembly was found to be 1:1 based on the integration of the picrate proton resonance at RT (Figure 3.4). Due to complications of symmetry, there is more than one structure consistent with this stoichiometry (e.g. isolated G-quartet, homodimer). Conclusive evidence for the identification of TASQ 10c·Na<sup>+</sup> came from diffusion NMR analysis.

**Figure 3.4.** Picrate signal integration for TASQ **10c**·Na<sup>+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.



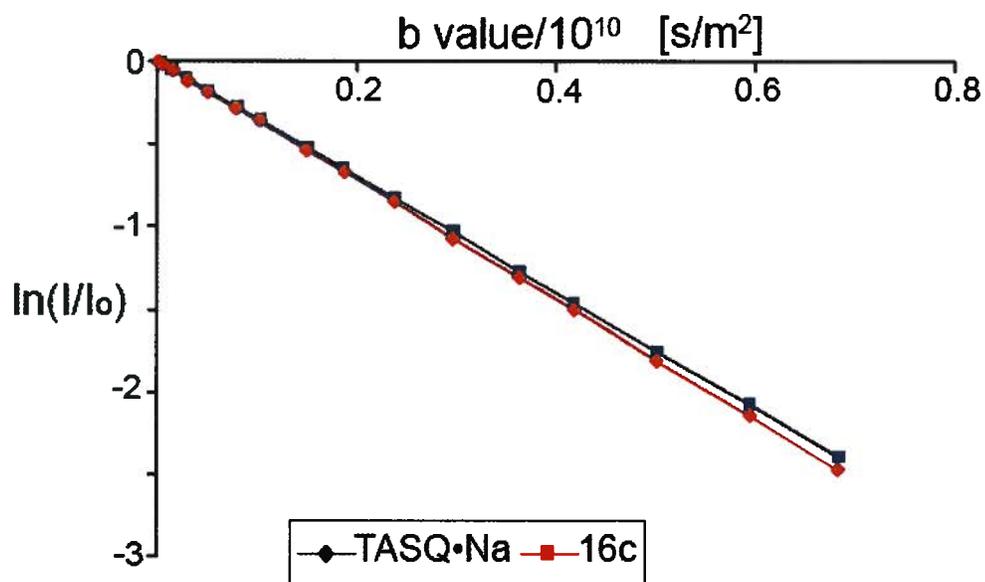
### 3.3.3.1 The Size Determination of TASQ **10c**·Na<sup>+</sup> by PFG NMR Spectroscopy

Initially, we attempted to use TASQ **10c** as an internal standard for TASQ **10c**-cations, but <sup>1</sup>H NMR experiments revealed that, with the exception of TASQ **10c**·Cs<sup>+</sup>, the assemblies undergo a rapid exchange with the cation-free TASQ, and give an average spectrum at room temperature. Therefore, diffusion measurements of TASQ **10c**·Na<sup>+</sup> were carried out in the presence of compound **16c** (the adenine analog of **10c**, Scheme 2.5). The experimental conditions used for the analysis were similar to those described for the cation-free TASQ in the section 2.13.11.3. The results are summarized in Table 3.1 and Figures 3.5 and 3.6.

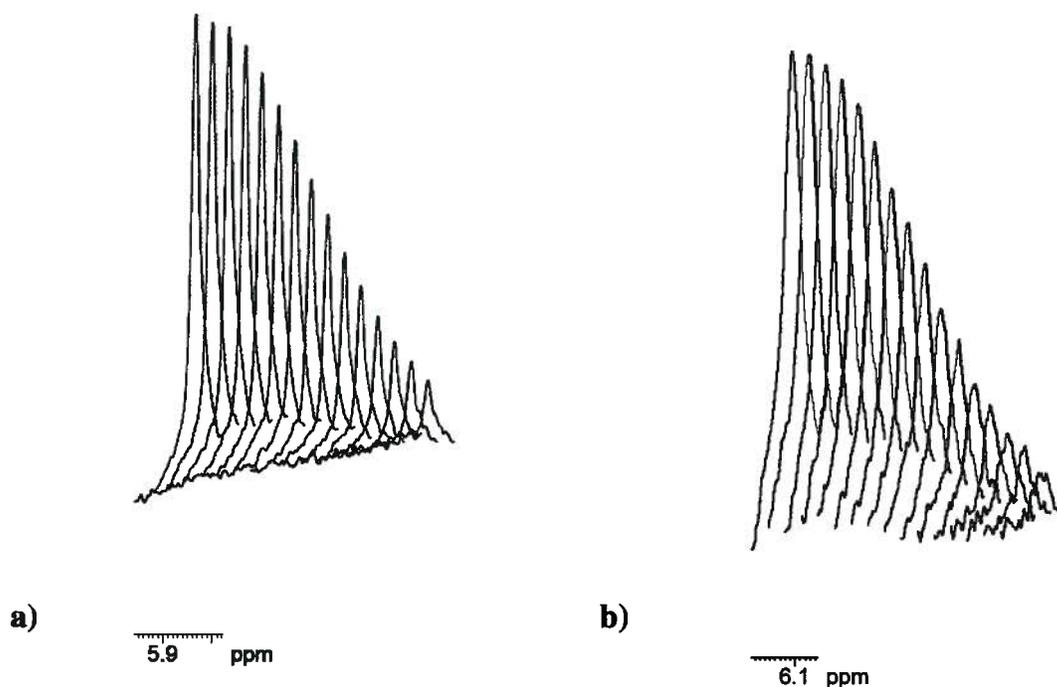
**Table 3.1.** Diffusion coefficients (D) of **10c**·Na<sup>+</sup> and **16c** in CDCl<sub>3</sub> at 295 K.

Solvent	D <sub>10c</sub> [×10 <sup>-10</sup> m <sup>2</sup> /s]	D <sub>10c·Na<sup>+</sup></sub> [×10 <sup>-10</sup> m <sup>2</sup> /s]	% Ratio
CDCl <sub>3</sub>	3.60 ± 0.10	3.48 ± 0.10	97

**Figure 3.5.** Normalized signal decay as a function of the b value at 295 K for **10c·Na<sup>+</sup>-16c** in CDCl<sub>3</sub>. The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value =  $(2 \cdot \pi \cdot \gamma \cdot g \cdot \delta)^2 (\Delta \cdot \delta / 3) \text{ s/m}^2$ ]



**Figure 3.6.** Normalized signal decay as a function of the gradient strength (g) at 295 K for a) the H1' signal of **10c·Na<sup>+</sup>**, and b) the H1' signal of **16c** in CDCl<sub>3</sub>.



The data listed in Table 3.1 suggest that TASQ **10c**·Na<sup>+</sup> forms a monomeric assembly (Figure 3.5). These results agree well with the stoichiometries obtained from picrate signal integration. To the best of our knowledge this is the first isolated cation-bound unimolecular G-quartet reported in solution-phase: It is known that the coordination of cations to the carbonyl oxygens of the guanines is the dominant force promoting the stacking of the bases in lipophilic systems.<sup>30</sup> Therefore, lipophilic G-quartets tend to stack such that the self-assembly process is very unlikely to stop at the quartet level in the presence of cations.<sup>31</sup> Indeed, the presence of cation-templated isolated G-quartets has only been reported in gas-phase.<sup>32</sup>

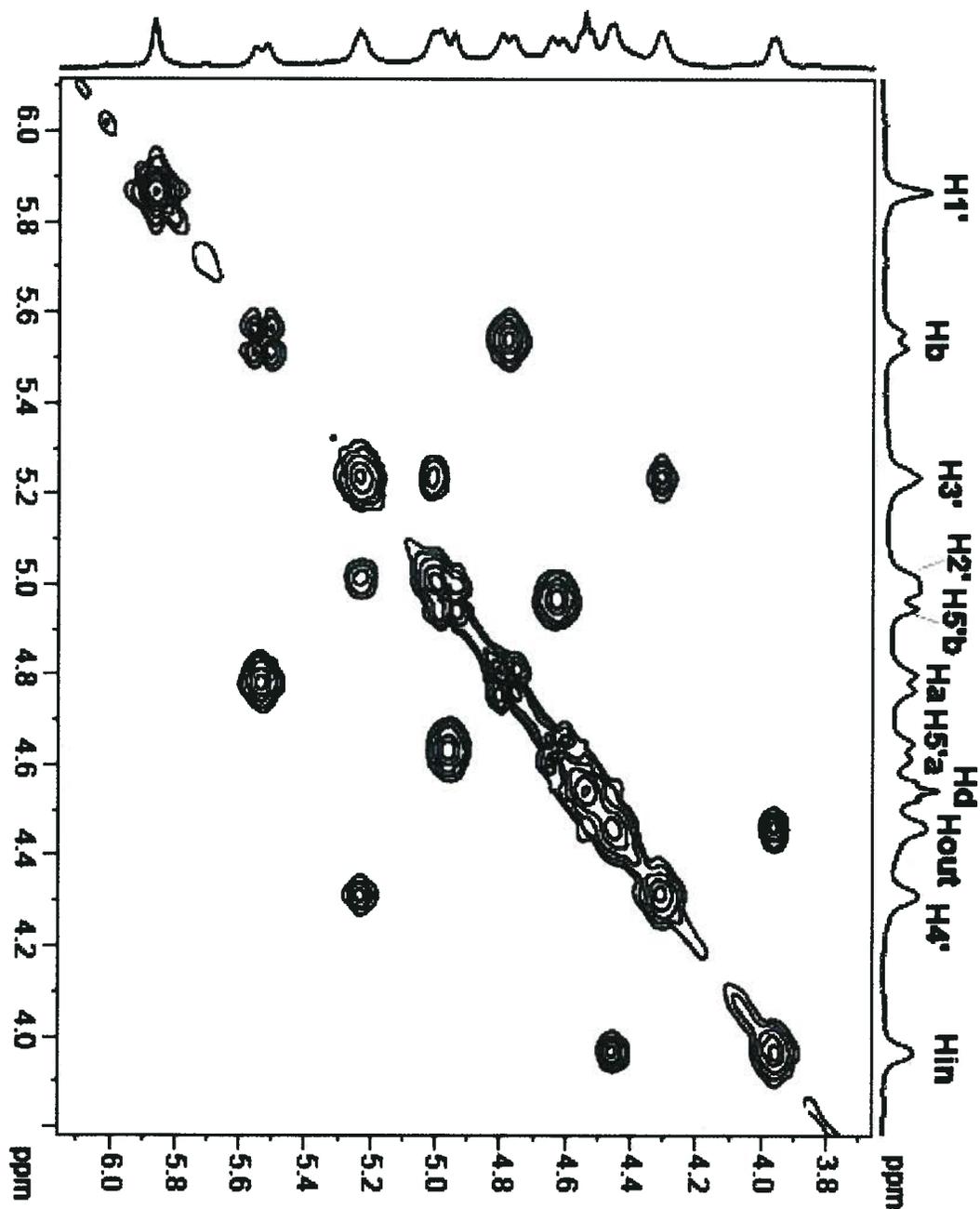
### 3.3.4 TASQ **10c**·Sr<sup>2+</sup>

Sr<sup>2+</sup> is a divalent cation which has been shown to promote formation of both lipophilic and hydrophilic G-quadruplexes.<sup>33-37</sup> The ionic radius of Sr<sup>2+</sup> is larger than that of Na<sup>+</sup> but smaller than that of K<sup>+</sup>.<sup>1</sup> TASQs show affinity for Sr<sup>2+</sup> and forms complexes whose NMR spectra are characterized by the following features:

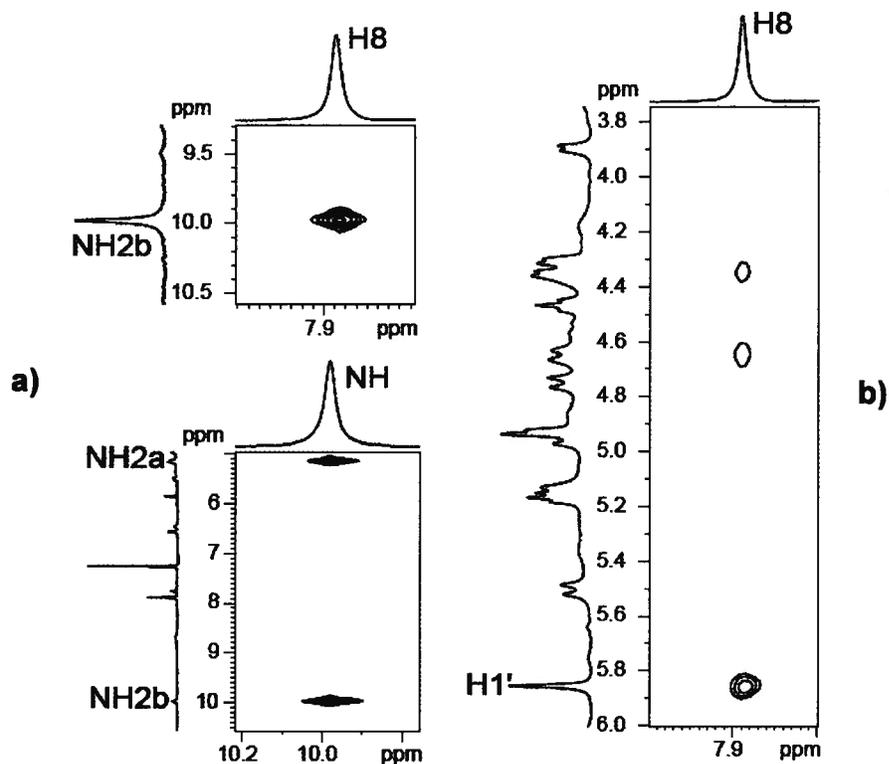
1) The presence of only one set of <sup>1</sup>H NMR signals which is indicative of the C<sub>4</sub> symmetry of the system (Figures 3.7); 2) The upfield shift of the H<sub>in</sub> proton which is indicative of the cation-binding (Figure 3.7); 3) The NOEs between H1'/H8 (Figure 3.8b), and H8/NH<sub>2b</sub> (Figure 3.8a) which are indicative of the *syn* conformation and the formation of G-quartet, respectively, 4) The 2:1 molar ratio of picrate per quartet which is indicative of the coordination of only one Sr<sup>2+</sup> per quartet (Figure 3.9). In addition to these features, the more pronounced upfield shift of the imino (NH) signal in the <sup>1</sup>H NMR spectrum of TASQ **10c**·Sr<sup>2+</sup> (compared to

those of TASQ **10c** and TASQ **10c**·Na<sup>+</sup> ) is most likely due to the interaction with the picrate anion which will be discussed in the following sections (Figure 3.1).

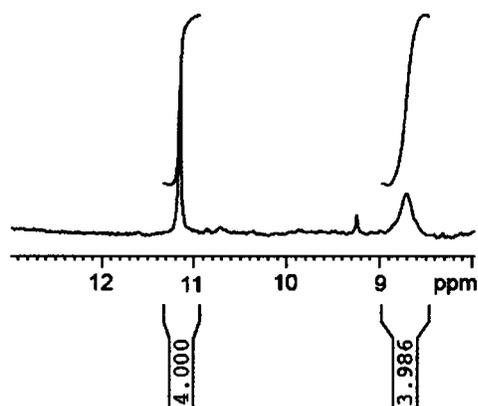
**Figure 3.7.** <sup>1</sup>H-<sup>1</sup>H COSY of TASQ **10c**·Sr<sup>2+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.



**Figure 3.8.** Portions of a 100 ms NOESY spectrum of TASQ 10c·Sr<sup>2+</sup> acquired at 400 MHz at -40 °C in CDCl<sub>3</sub> indicative of a) the formation of G-quartet, and b) the *syn* conformation. The NH<sub>2</sub> signals which are broadened up at RT are visible at low temperature.



**Figure 3.9.** Picrate signal integration for TASQ 10c·Sr<sup>2+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.



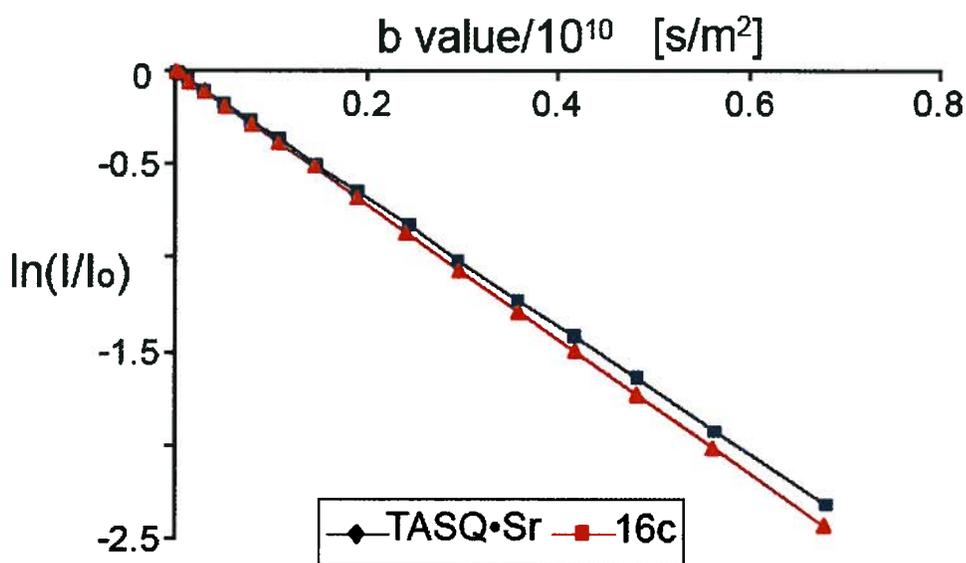
### 3.3.4.1 The Size Determination of TASQ 10c·Sr<sup>2+</sup> by PFG NMR Spectroscopy

The diffusion NMR experiments of TASQ 10c·Sr<sup>2+</sup> were conducted in a way similar to TASQ 10c·Na<sup>+</sup>. An equimolar mixture of 16c-10c·Sr<sup>2+</sup> in CDCl<sub>3</sub> was prepared and subjected to PFG NMR spectroscopic analysis. The results are summarized in Table 3.2 and Figures 3.10 and 3.11.

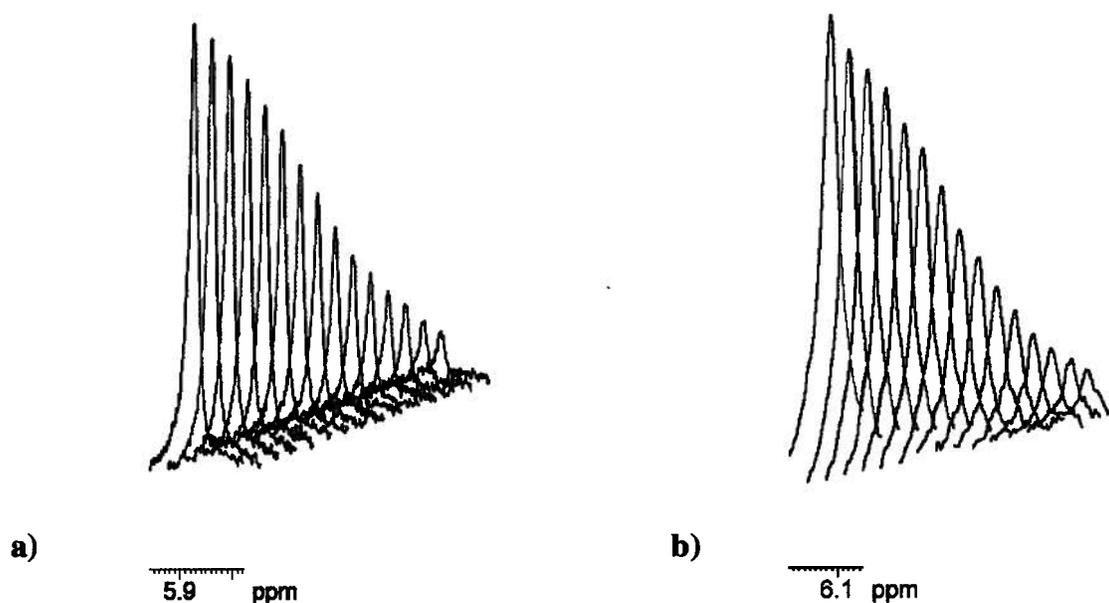
**Table 3.2.** Diffusion coefficients (D) of 10c·Sr<sup>2+</sup> and 16c in CDCl<sub>3</sub> at 295 K.

Solvent	D <sub>10c</sub> [×10 <sup>-10</sup> m <sup>2</sup> /s]	D <sub>10c·Sr<sup>2+</sup></sub> [×10 <sup>-10</sup> m <sup>2</sup> /s]	% Ratio
CDCl <sub>3</sub>	3.56 ± 0.10	3.41 ± 0.10	96

**Figure 3.10.** Normalized signal decay as a function of the b value at 295 K for 10c·Sr<sup>2+</sup>-16c in CDCl<sub>3</sub>. The diffusion coefficients (D) were calculated for the representative signal of H1'. [b value = (2·π·γ·g·δ)<sup>2</sup>(Δ-δ/3) s/m<sup>2</sup>]



**Figure 3.11.** Normalized signal decay as a function of the gradient strength ( $g$ ) at 295 K for a) the H1' signal of  $10c \cdot Sr^{2+}$ , and b) the H1' signal of  $16c$  in  $CDCl_3$ .



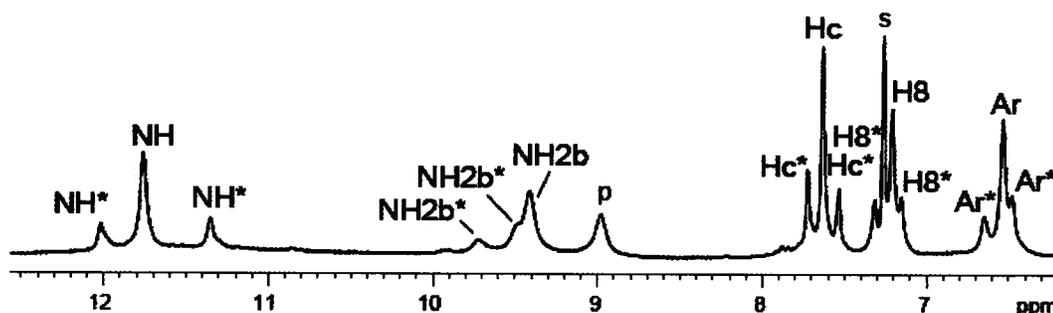
The results of diffusion NMR analysis suggest that TASQ  $10c \cdot Sr^{2+}$  forms an isolated G-quartet. This indicates that TASQ  $10c \cdot Sr^{2+}$  and TASQ  $10c \cdot Na^+$  are not only structurally, but also stoichiometrically, similar to each other. The cation in both complexes seems to be coordinated by only four guanines and forms no stacked layers. The possible reasons for the formation of isolated G-quartets, including the role of anion, will be discussed in the following sections.

### 3.3.5 TASQ $10c \cdot K^+$

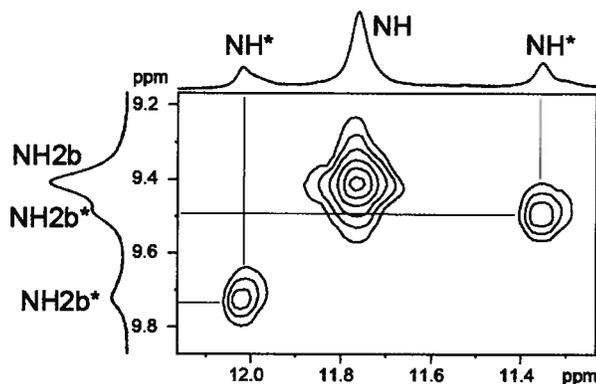
Potassium is one of the most preferred cations for the formation of nucleic acid G-quadruplexes.<sup>1,38</sup> TASQ  $10c$  binds potassium and gives a broad spectrum at room temperature (Figure 3.1). At low temperature the spectrum splits into three sets of signals indicating a

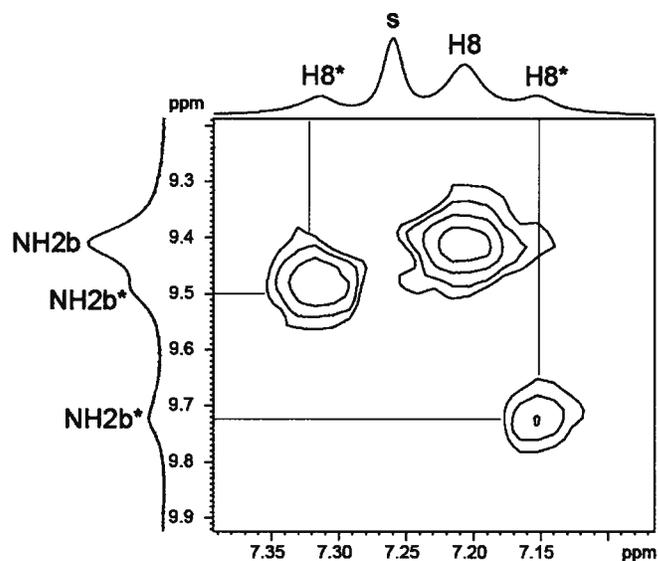
dynamic mixture of structures (Figure 3.12). It can be seen from Figure 3.13 that all the signals show the characteristic NOE correlations of G-quartet assemblies (NOEs between H8/NH<sub>2b</sub> and NH/NH<sub>2</sub>). The main set of signals shows a strong H8/H1' correlation and appears to be originating from an all-*syn*<sup>39,40</sup> G-quartet similar to those of Na<sup>+</sup> and Sr<sup>2+</sup>. The side signals in Figure 3.12 have equal intensities and seem to be related to a second “new” species present in the solution. From signal integration, it can be approximated that this second species comprises almost one-third of the population at -40 °C (Figure 3.14).

**Figure 3.12.** Partial <sup>1</sup>H NMR spectrum of TASQ 10c·K<sup>+</sup> at 400 MHz at -40 °C displaying three sets of signals. The two sets of small signals related to a second “new” species have been marked with asterisks. The signals labeled “p” and “s” corresponds to the picrate and the solvent, respectively.

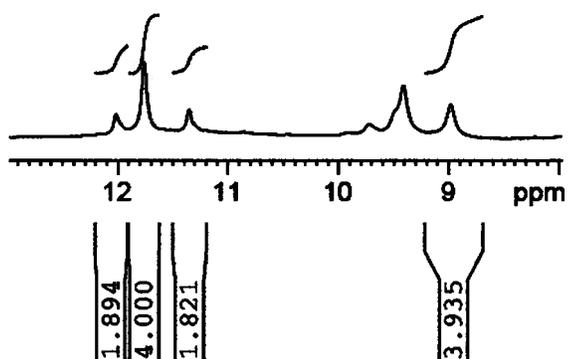


**Figure 3.13.** Portions of a 100 ms NOESY spectrum of TASQ 10c·K<sup>+</sup> acquired at 400 MHz in CDCl<sub>3</sub>.





**Figure 3.14.** Picrate signal integration for TASQ 10c·K<sup>+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.

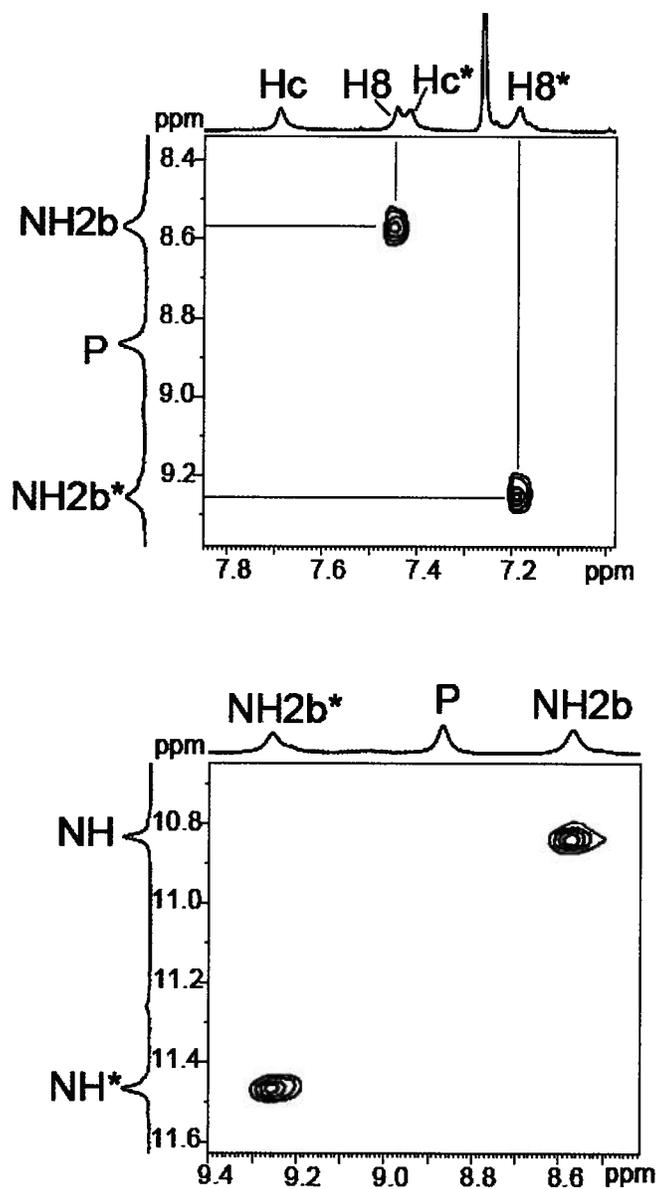


The information regarding the stoichiometry of TASQ 10c·K<sup>+</sup> came from the analysis of TASQ 10c·Cs<sup>+</sup> (section 3.3.6.1). Only the two “new” sets of signals are observed in the presence of the large Cs<sup>+</sup>. This reveals the presence of two different modes of binding for TASQs that depends on the size of the cations. Potassium is apparently in the intermediate size-range, resulting in a mixture in which the two structures have comparable stability.

### 3.3.6 TASQ 10c·Cs<sup>+</sup>

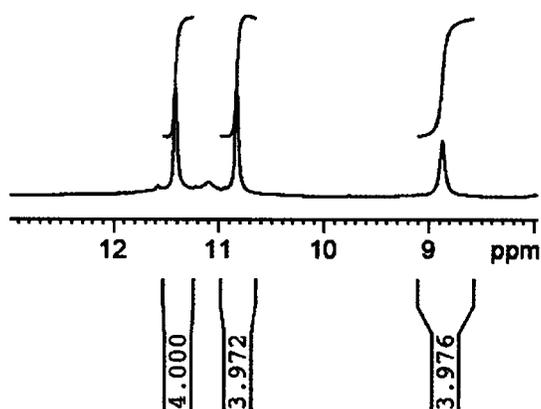
Cs<sup>+</sup> is probably the least studied cation among the four selected cations. TASQs bind cesium and display two sets of <sup>1</sup>H NMR signals. 2D NOESY experiments confirm that the cesium-bound complexes retain the G-quartet structure (Figure 3.15).

**Figure 3.15.** Partial <sup>1</sup>H-<sup>1</sup>H NOESY of TASQ 10a·Cs<sup>+</sup> acquired at 400 MHz at -40 °C in CDCl<sub>3</sub>. Signals shown are indicative of the formation of G-quartet.



Based on the picrate proton resonance integration at RT, the ratio of cation per quartet was found to be 1:1 for TASQ 10c·Cs<sup>+</sup> (Figure 3.16). This ratio is identical to the ratios found for the assemblies of Na<sup>+</sup> and Sr<sup>2+</sup>. However, the symmetry of the TASQs is clearly broken in cesium (and potassium). The unambiguous stoichiometry of TASQ 10c·Cs<sup>+</sup> was obtained from diffusion NMR experiments.

**Figure 3.16.** Picrate signal integration for TASQ 10c·Cs<sup>+</sup> at 400 MHz in CDCl<sub>3</sub> at 25 °C.



### 3.3.6.1 The Size Determination of TASQ 10c·Cs<sup>+</sup> by PFG NMR Spectroscopy

The TASQ·cation assemblies which have been analyzed by the PFG method so far have all turned out to be monomeric. Before providing the results of the analysis for TASQ 10c·Cs<sup>+</sup>, it is necessary to describe the relationship of the diffusion coefficients (*D*) of a monomer and a dimer. From Chapter 2 (section 2.3.11.1), we know that the diffusion coefficient of a molecule (*D*) is related to its hydrodynamic radius (*R*) by Equation 1:

$$D = \frac{k_b \cdot T}{6 \cdot \pi \cdot \eta \cdot R} \quad (1)$$

Thus, the relationship between the diffusion coefficient of a monomer and a dimer will be given by Equation 2:

$$\frac{D_d}{D_m} = \frac{R_m}{R_d} \quad (2)$$

Assuming that both monomer and dimer are spherical, the hydrodynamic radii of both molecules can be calculated from their corresponding hydrodynamic volumes (Equations 3-6)

$$V = \frac{4}{3} \pi \cdot R^3 \quad (3)$$

$$R_m = \sqrt[3]{\frac{3 \cdot V_m}{4 \cdot \pi}} \quad (4)$$

$$R_d = \sqrt[3]{\frac{3 \cdot (2 \cdot V_m)}{4 \cdot \pi}} = \sqrt[3]{2} \cdot \sqrt[3]{\frac{3 \cdot V_m}{4 \cdot \pi}} \quad (5)$$

By replacing the Equations (4) and (5) into the Equation (2)

$$\frac{D_d}{D_m} = \frac{1}{\sqrt[3]{2}} = 0.79$$

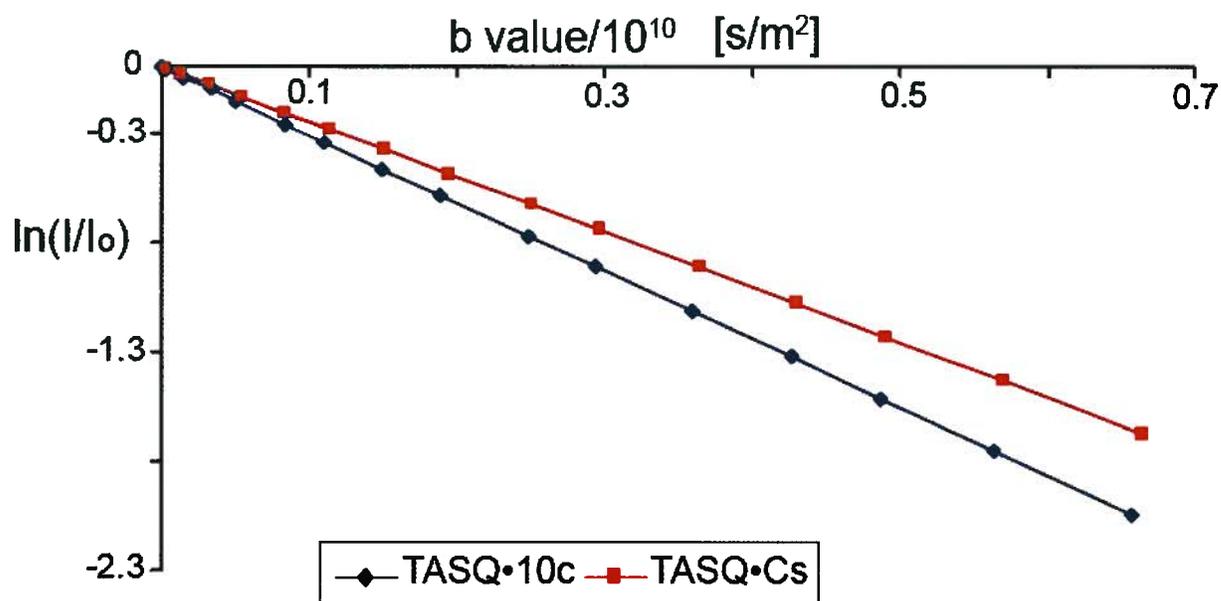
Thus, the diffusion coefficient of a dimer is approximately 79% of that of a monomer.

The diffusion of  $10\mathbf{c}\cdot\text{Cs}^+$  was directly measured in the presence of free  $10\mathbf{c}$  for two reasons. First,  $10\mathbf{c}$  and  $10\mathbf{c}\cdot\text{Cs}^+$  undergo sufficiently slow exchange at RT so that they could be analyzed in the same solution. Second, the use of  $16\mathbf{c}$  as internal standard was problematic due to overlapping of signals. Thus, an equimolar mixture of  $10\mathbf{c}\cdot\text{Cs}^+$ - $10\mathbf{c}$  was prepared and subjected to analysis. The results are summarized in Table 3.3 and Figures 3.17 and 3.18.

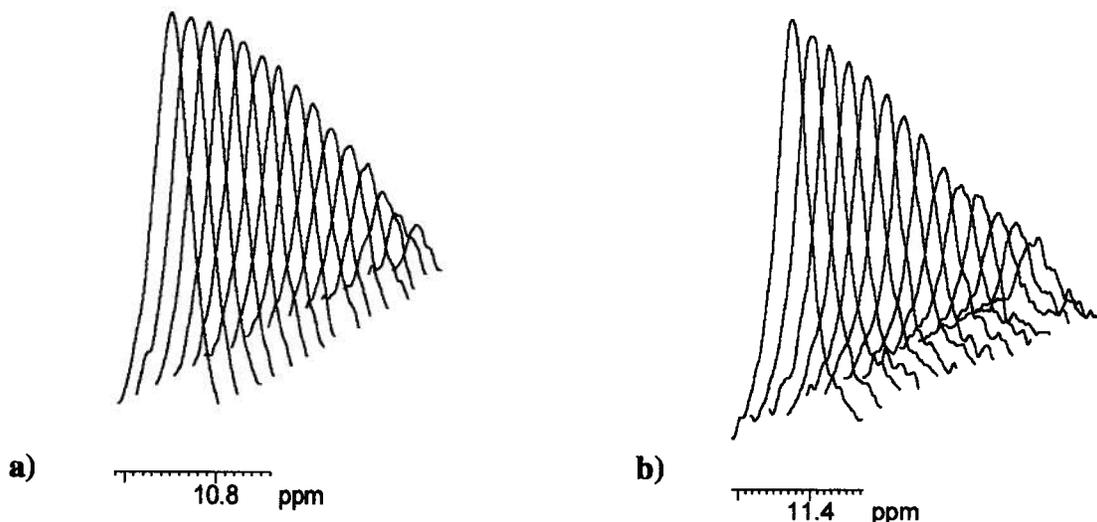
**Table 3.3.** Diffusion coefficients (D) of  $10\mathbf{c}\cdot\text{Cs}^+$  and  $10\mathbf{c}$  in  $\text{CDCl}_3$  at 295 K.

Solvent	$D_{10\mathbf{c}}$ [ $\times 10^{-10} \text{ m}^2/\text{s}$ ]	$D_{10\mathbf{c}\cdot\text{Cs}^+}$ [ $\times 10^{-10} \text{ m}^2/\text{s}$ ]	% Ratio
$\text{CDCl}_3$	$3.11 \pm 0.10$	$2.50 \pm 0.10$	80

**Figure 3.17.** Normalized signal decay as a function of the b value at 295 K for  $10\mathbf{c}\cdot\text{Cs}^+$ - $10\mathbf{c}$  in  $\text{CDCl}_3$ . The diffusion coefficients (D) were calculated for the representative signal of NH. [b value =  $(2\cdot\pi\cdot\gamma\cdot g\cdot\delta)^2(\Delta\cdot\delta/3) \text{ s}/\text{m}^2$ ]



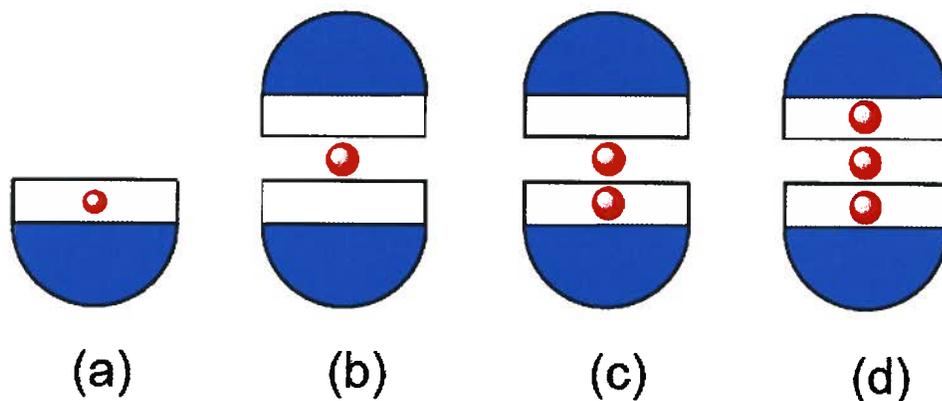
**Figure 3.18.** Normalized signal decay as a function of the gradient strength ( $g$ ) at 295 K for a) the NH signal of  $10\mathbf{c}\cdot\text{Cs}^+$ , and b) the NH signal of  $10\mathbf{c}$  in  $\text{CDCl}_3$ .



The diffusion NMR experiments show that in contrast to the behavior of small cations, cesium promotes the formation of a dimer. The experimental ratio of 80% obtained for  $D_{10\mathbf{c}\cdot\text{Cs}^+}/D_{10\mathbf{c}}$  is in close agreement with the theoretical value of 79% calculated for a dimeric system. Considering the asymmetry found in the  $^1\text{H}$  NMR spectra, the 1:1 ratio of cation per quartet, and the symmetry of the system, this dimer has to be an asymmetric heterodimer such as structure (c) in Figure 3.19. The coordination of a third cation is probably prevented due to repulsive forces between cations.

The results obtained for  $10\mathbf{c}\cdot\text{Cs}^+$  also shed light on the identity of TASQ  $10\mathbf{c}\cdot\text{K}^+$  and suggest that it is most likely a mixture of a monomer and a dimer such as (a) and (c) of Figure 3.19, which exchange rapidly at room temperature. The next section will outline the role of the anion in the formation and stabilization of these structures.

**Figure 3.19.** Proposed G-quartet-based structures of TASQ-cation assemblies a) isolated G-quartet, b,d) homodimers, and c) heterodimer. Only structures consistent with (a) and (c) were observed.



### 3.3.7 The Anion Effect

Picrate salts have extensive applications in extraction experiments, and their interactions with various ionophores such as crown ethers have been investigated.<sup>41,42</sup> They are only slightly soluble in chloroform ( $\epsilon_r = 4.8$ ), and exist mostly as ion pair complexes in this media.<sup>43</sup> The picrate anion is a mono or bidentate ligand which can coordinate to alkaline or alkaline-earth metal cations through its phenolic and *o*-nitro groups.<sup>42</sup> It can be seen from Figure 3.1 that the picrate anion interacts with the TASQ-cation assemblies. The upfield shift of the picrate signal in the monomeric assemblies suggests that the picrate anion is associated with the cations and is positioned near or above the aromatic guanines.<sup>44</sup> Similar observations have been reported by Davis et al. who demonstrated that a capping picrate anion can stabilize a sodium-templated octamer and prevent it from converting to a hexadecamer.<sup>45</sup> We found that the capping effect of the anion in TASQs is controlled by the size of the cation and is more pronounced for the smaller

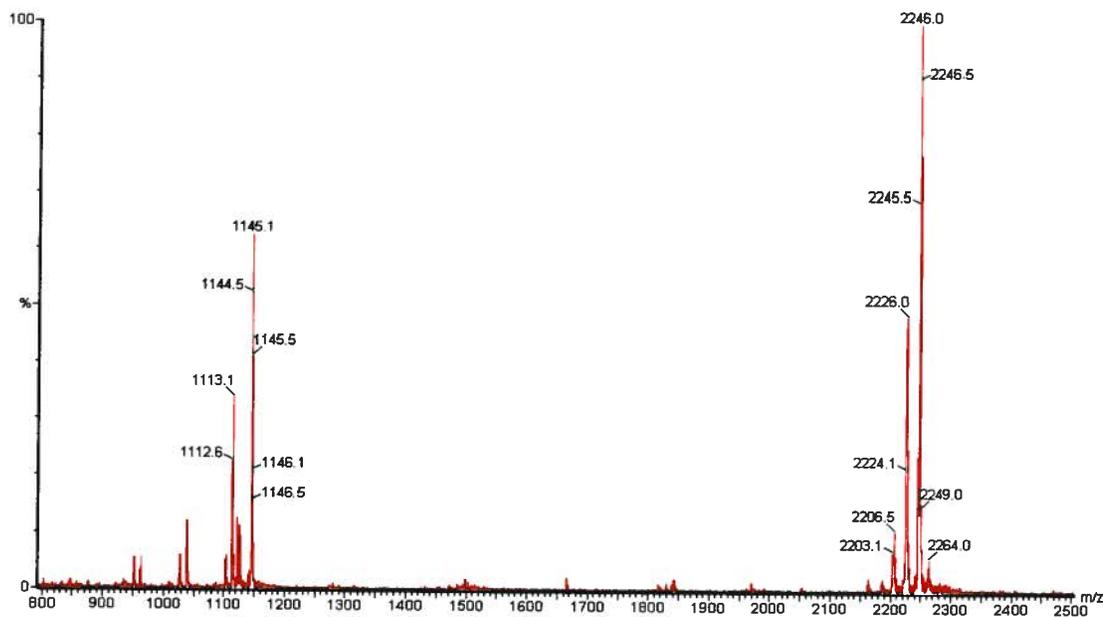
cations. Larger cations such as cesium (and potassium somewhat) prefer formation of dimeric systems in which ion pairs are completely dissociated.

A set of extraction experiments in the presence of non-coordinating anions ( $\text{I}^-$ ,  $\text{BPh}_4^-$ ) verified our conclusion. Under these new conditions,  $\text{TASQ } 10\text{c}\cdot\text{Cs}^+$  gave a spectrum similar to the one recorded in the presence of picrate anion. This suggests that the formation of heterodimer is not affected by the anion.  $\text{TASQ } 10\text{c}\cdot\text{Na}^+$  and  $\text{TASQ } 10\text{c}\cdot\text{Sr}^{2+}$ , by contrast, yielded very broad spectra indicating the loss of the monomer as the singular species (data not shown). Effort is underway to characterize the newly formed structures, which are likely a mixture of higher order assemblies.

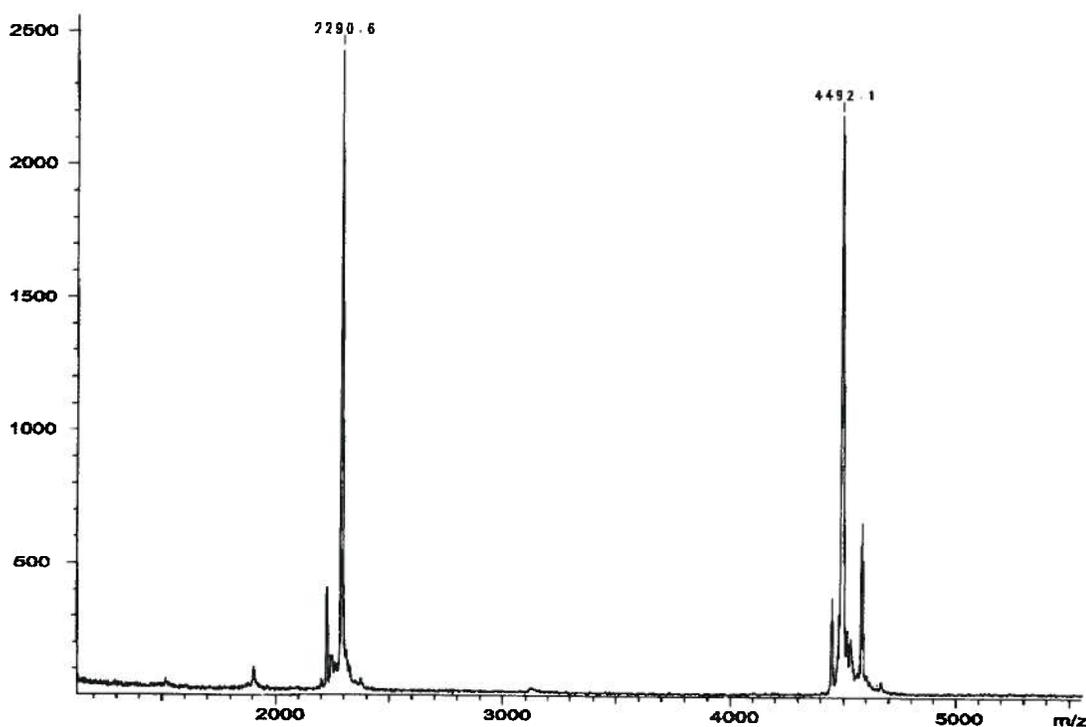
### 3.3.8 Gas-Phase Structures

As was outlined in the preceding sections, the bulk of the experiments employed for the characterization of  $\text{TASQ}\cdot\text{cation}$  assemblies were NMR experiments. With the help of NMR methods, we could identify the G-quartet as the main structural units of the  $\text{TASQ}\cdot\text{cation}$  assemblies and could discriminate between the monomeric and dimeric systems. We could also demonstrate the influence of the anions on the morphology of  $\text{TASQ}\cdot\text{cation}$  assemblies. In a complementary set of experiments, mass spectrometric methods were used to analyze a monomeric system. We were mainly interested in perturbing the interaction of the anion with the metal center in a monomeric  $\text{TASQ}$  and monitoring the resulting changes in a relatively quick and easy experiment. The monomeric assembly of  $\text{TASQ } 10\text{a}\cdot\text{Sr}^{2+}$  was selected for this purpose and subjected to MALDI-MS and ESI-MS analysis. The results are shown in Figures 3.20 and 3.21.

**Figure 3.20.** ESI-MS spectrum of TASQ 10a·Sr<sup>2+</sup> in positive mode. Two species of [TASQ 10a·Sr]<sup>2+</sup> (m/z = 1145) and [(TASQ 10a)<sub>2</sub>·Sr]<sup>2+</sup> (m/z = 2246) were observed in the gas-phase.



**Figure 3.21.** MALDI-MS spectrum of TASQ 10a·Sr<sup>2+</sup> in positive mode. Two species of [TASQ 10a·Sr]<sup>+</sup> (m/z = 2290) and [(TASQ 10a)<sub>2</sub>·Sr]<sup>+</sup> (m/z = 4492) were observed in the gas-phase (note that MALDI typically generates only singly charged species)



These results indicate the presence of a monomer (corresponding to an isolated G-quartet) and a dimer (corresponding to a homodimer) in the gas-phase. The signals of sodium-bound complexes are most likely the result of cation-exchange during the analysis. The presence of monomeric and dimeric species, and the low intensity of signals related to polymeric aggregates suggest that: 1) the observed structures are the results of cation-coordination and not the random H-bonding of the TASQ molecules, and 2) the observed structures retain the G-quartet motif by organized coordination to the cations.

As is illustrated in Figures 3.20 and 3.21, peaks related to the association of the picrate anion to any of the clusters were not observed in the recorded spectra. This suggests the high lability of the picrate anion under the experimental conditions and also provides an explanation for the formation of the dimeric assembly.

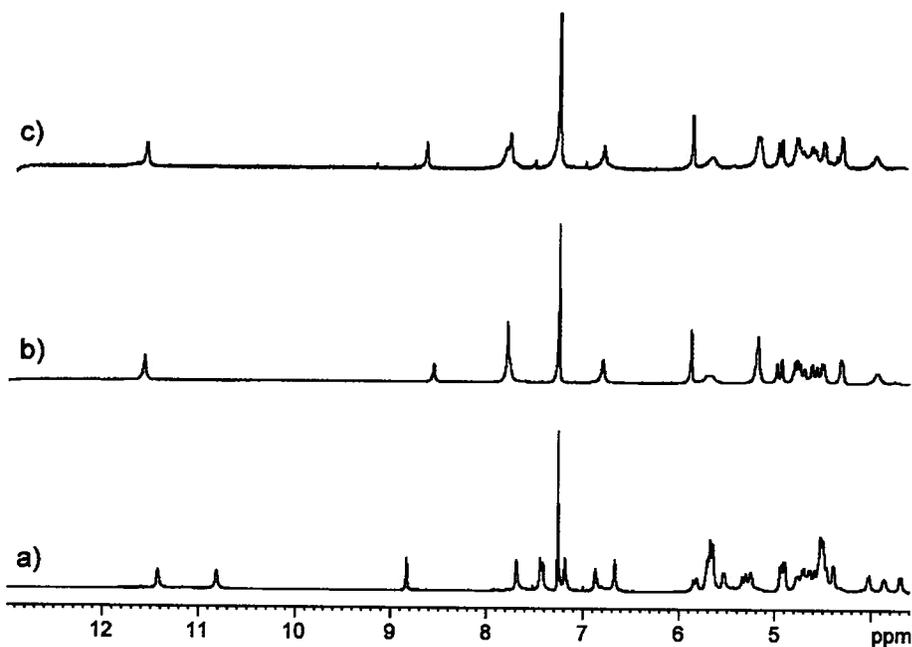
### 3.3.9 Selectivities of TASQs

Cation-induced stabilization of nucleic acid G-quadruplexes is governed by two important factors, i) free energy of cation-binding and ii) heat of dehydration.<sup>1,46,47</sup> The latter factor is of less importance for lipophilic G-quadruplexes. TASQs show affinity for a broad range of cations. Among those, the facile extraction of cesium came as a surprise. Simple guanosine compounds such as guanosine 5'-monophosphate (5'-GMP) generally demonstrate very low affinity for cesium.<sup>7</sup> The large ionic radius of cesium results in substantial separation of quartets (~4.6 Å) which is not favorable for the optimal stacking of the bases (~3.3 Å).<sup>21</sup> Therefore, few experiments have been carried out on cesium-binding by guanine assemblies. This contrasts with the many studies of Na<sup>+</sup> and K<sup>+</sup> complexation by guanine assemblies.

TASQs show greater cesium affinity than their building blocks. The cesium affinity of compound **9** is likely to be comparable to that of 5'-GMP. The parent scaffold of TASQs, a cavitand with four phenolic groups has also been shown to have a very low affinity for cesium.<sup>48</sup> The formation of TASQs enhances the cesium affinity over its constituents due to the preorganization of guanines into a G-quartet. However, TASQs are not cesium-selective. Addition of solid sodium picrate to a solution of TASQ **10c**·Cs<sup>+</sup> in chloroform resulted in complete replacement of cesium ions with sodium. This process is also accompanied by a major structural change in which the dimeric system converts to a monomer. In contrast, addition of solid cesium picrate to a solution of TASQ **10c**·Na<sup>+</sup> in chloroform resulted in no change. Moreover, extraction experiments performed on TASQ **10a** with an equimolar mixture of solid sodium and cesium picrates resulted in exclusive formation of TASQ **10a**·Na<sup>+</sup>. These results indicate that TASQs have a high Na<sup>+</sup>/Cs<sup>+</sup> selectivity (Figure 3.22). In a similar set of experiments performed on the complexes of Sr<sup>2+</sup> and Na<sup>+</sup>, the divalent cation completely replaced the sodium ion. The selectivity of TASQ for potassium was also evaluated to be between sodium and cesium. The results are summarized in Table 3.4. The clean separation of strontium from mixtures of other ions by TASQs may be of industrial significance. <sup>90</sup>Sr<sup>2+</sup>, and <sup>137</sup>Cs<sup>+</sup> are both products of nuclear fission and their selective stripping from salted radioactive waste solutions has been the subject of considerable research in the past.<sup>49,50</sup> It should also be noted that Table 3.4 displays only the selectivities in the presence of picrate anion. The anion as we reported in the previous section changes the structure and may modulate the overall selectivities.

A series of control experiments in which **16c** was mixed with either sodium or cesium picrates showed no changes either in the color or in the <sup>1</sup>H NMR spectra. Thus, guanines are essential for the cation-affinity of TASQs. The overall selectivity is Sr<sup>2+</sup>>Na<sup>+</sup>>K<sup>+</sup>>Cs<sup>+</sup>.

**Figure 3.22.** a)  $^1\text{H}$  NMR spectrum of pure TASQ  $10\text{a}\cdot\text{Cs}^+$ , b)  $^1\text{H}$  NMR spectrum of pure TASQ  $10\text{a}\cdot\text{Na}^+$ , c)  $^1\text{H}$  NMR spectrum recorded after extraction of TASQ  $10\text{a}$  with an equimolar mixture of sodium and cesium picrates.



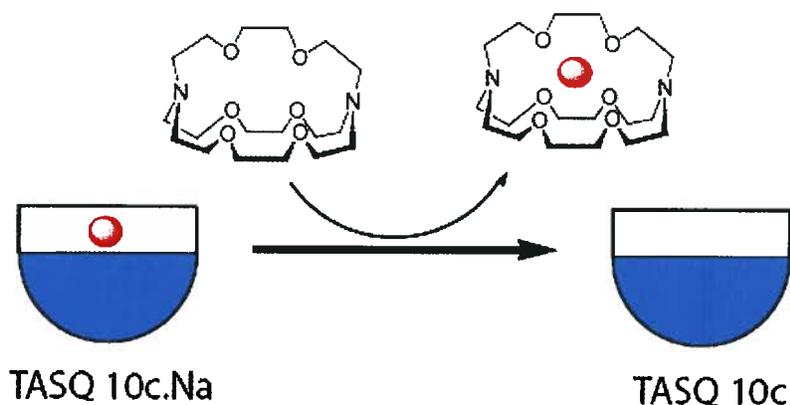
**Table 3.4.** The selectivities of TASQs. The magnitude of these selectivities is estimated as more than a factor of 20 between each pair of cations.

Starting Species	Added M·Picrate	Resulting Species
TASQ $10\text{c}\cdot\text{Na}^+$	Cs·Pic	TASQ $10\text{c}\cdot\text{Na}^+$
TASQ $10\text{c}\cdot\text{Cs}^+$	Na·Pic	TASQ $10\text{c}\cdot\text{Na}^+$
TASQ $10\text{a}$	Na·Pic + Cs·Pic	TASQ $10\text{a}\cdot\text{Na}^+$
TASQ $10\text{c}\cdot\text{Na}^+$	Sr·Pic	TASQ $10\text{c}\cdot\text{Sr}^{2+}$
TASQ $10\text{c}\cdot\text{Sr}^{2+}$	Na·Pic	TASQ $10\text{c}\cdot\text{Sr}^{2+}$
TASQ $10\text{c}\cdot\text{Na}^+$	K·Pic	TASQ $10\text{c}\cdot\text{Na}^+$
TASQ $10\text{c}\cdot\text{K}^+$	Na·Pic	TASQ $10\text{c}\cdot\text{Na}^+$
TASQ $10\text{c}\cdot\text{Cs}^+$	K·Pic	TASQ $10\text{c}\cdot\text{K}^+$
TASQ $10\text{c}\cdot\text{K}^+$	Cs·Pic	TASQ $10\text{c}\cdot\text{K}^+$

### 3.3.10 Interconversion of TASQ 10c·Na<sup>+</sup> to TASQ 10c

Cation-binding by TASQs is a reversible process. Up to this point, we have only focused on the affinity of TASQs for cations and the analysis of the resulting structures. One can expect that the sequestration of the cations with an appropriate ligand would result in the conversion of cation-bound TASQs to cation-free TASQs. Cryptands are a class of strong cation-binding agents which can be used for this purpose.<sup>51</sup> Lehn and Spada have independently shown that addition of [2.2.2] cryptand to G-quadruplex solutions causes a shift towards the formation of non G-quartet based systems (see Chapter 1).<sup>52,53</sup> TASQ is unique from this point of view, since the addition of cryptand to a cation-bound system is expected to shift the equilibrium towards the formation of another G-quartet based system (Figure 3.23).

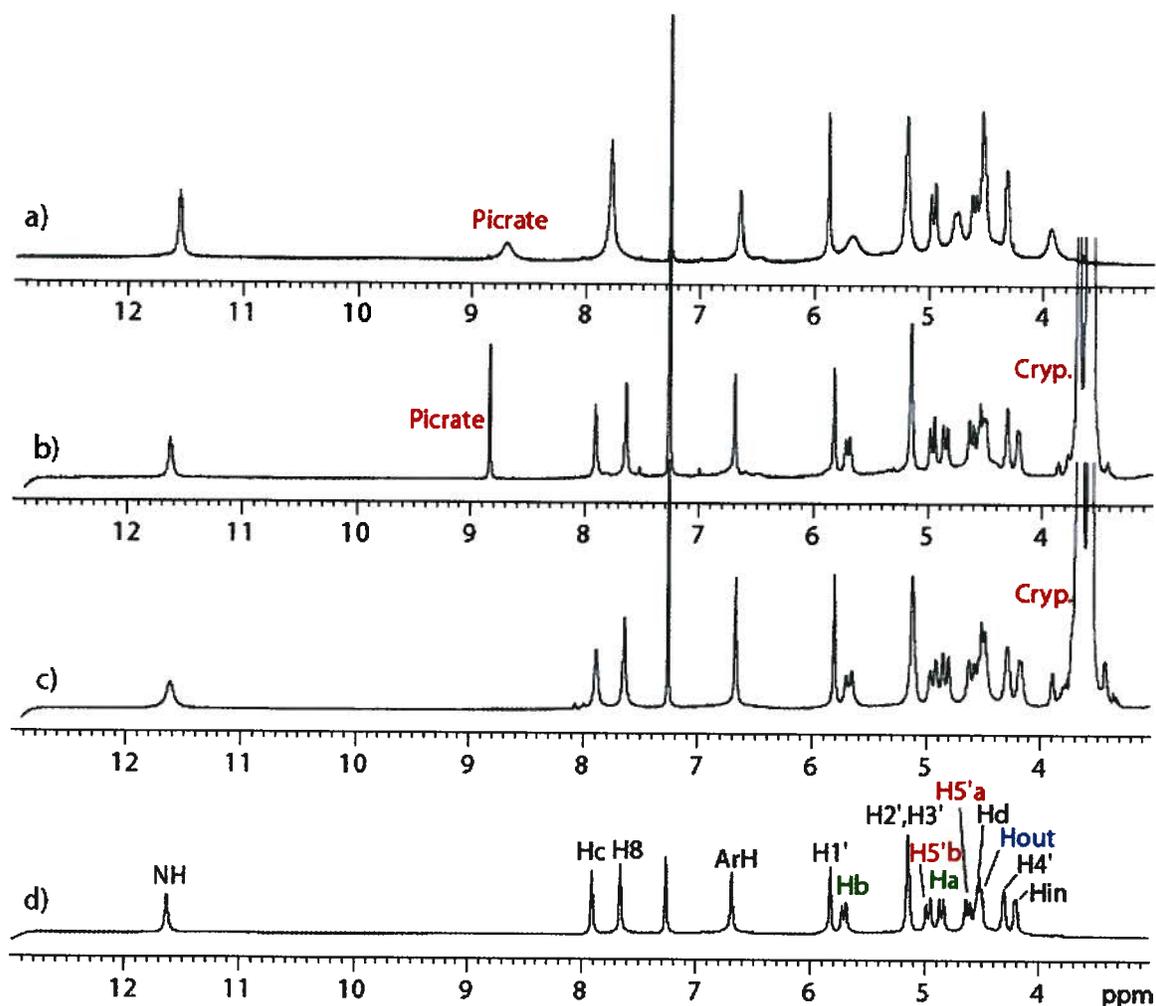
**Figure 3.23.** Schematic representation of the conversion of TASQ 10c·Na<sup>+</sup> to TASQ 10c upon addition of [2.2.2] cryptand.



As is shown in figure 3.24, addition of cryptand [2.2.2] to the solution of TASQ 10c·Na<sup>+</sup> resulted in an <sup>1</sup>H NMR spectrum identical to the spectrum of the cation-free species. Addition of

cryptand [2.2.2] to the cation-free TASQ resulted in no change. This latter experiment confirms our findings that TASQ **10c** is indeed a cation-free species. Note that in addition to the distinction made here, the atomic absorption analyses (section 3.5.4) further distinguish TASQ **10c**·Na<sup>+</sup> from TASQ **10c**.

**Figure 3.24.** a) <sup>1</sup>H NMR spectrum of a 2.4 mM solution of TASQ **10c**·Na<sup>+</sup> at 400 MHz in CDCl<sub>3</sub> at 25°C, b) The spectrum recorded after addition of the 8-fold excess of [2.2.2] cryptand to the same solution (note the free picrate signal at ~8.8 ppm), c) <sup>1</sup>H NMR spectrum of a 2.4 mM of TASQ **10c** after addition of the 8-fold excess of [2.2.2] cryptand, and d) <sup>1</sup>H NMR spectrum of a solution of TASQ **10c** at 400 MHz in CDCl<sub>3</sub> at 25°C. (Note: attempts for the conversion of TASQ **10c**·Na<sup>+</sup> to TASQ **10c** by adding 5-crown-5 was unsuccessful due to the higher association constant of TASQ with Na<sup>+</sup>).



### 3.3.11 CD Spectroscopy

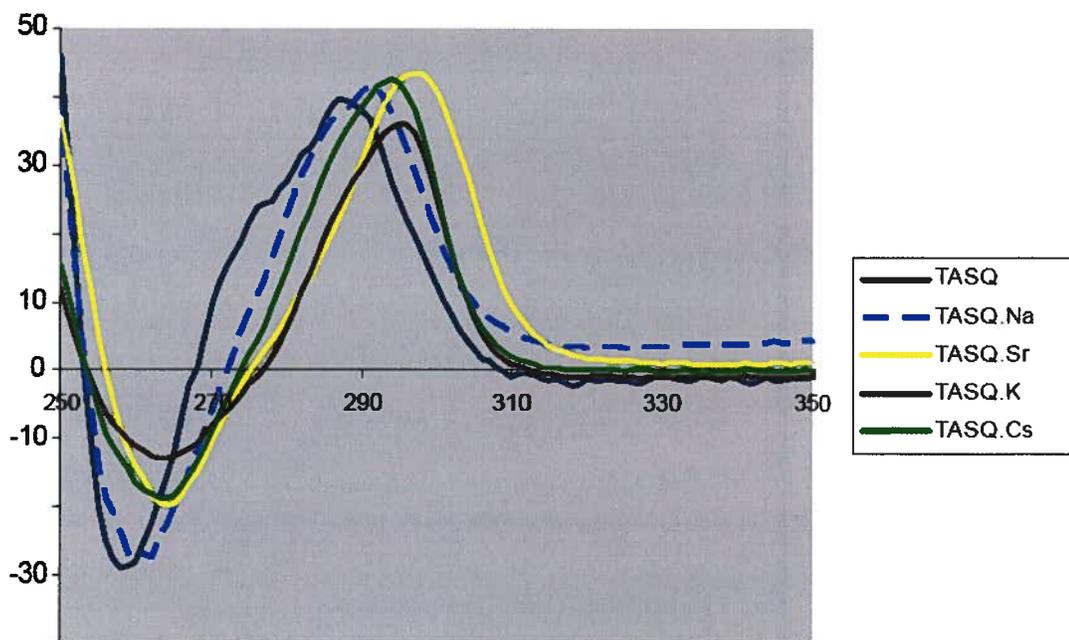
Circular dichroism (CD) spectra are used extensively for studying the structure of lipophilic and nucleic acid G-quadruplexes.<sup>54-56</sup> An exciton couplet centered at ~270 nm (guanine absorption region) is generally considered to be the sign of the formation of an aggregate composed of at least two stacking G-quartets.<sup>22</sup> Some information such as distortion of G-quartet layers can be deduced from the shape of the spectra. Moreover, the polarity of CD spectra gives some information about the clockwise or counterclockwise rotation of G-quartet layers with respect to each other.<sup>54</sup>

Due to structural differences, the information obtained from the study of lipophilic or nucleic acid G-quadruplexes cannot be directly applied to TASQ. In contrast to the above mentioned systems, TASQ is a singular G-quartet in which all the building blocks are parts of the same molecule. Moreover, the CD spectrum of TASQ is affected significantly by the presence of other chromophores. The cavitand and triazole components of TASQs both contain aromatic moieties whose  $\lambda_{\text{max}}$ 's are similar to that of guanine itself. Therefore, the observed CD spectrum can be considered the result of the interaction of all chromophores. This poses an obstacle for the detailed interpretation of the CD spectra recorded for TASQs.

The CD spectra of TASQ-cation assemblies exhibit a negative and a positive cotton effect for all of the complexes in chloroform (Figure 3.25). These CD bands are all centered at around 270 nm and are most likely the results of similar interactions between chromophores. The comparison of the CD spectra shows that the shape and polarity of the spectra is independent of the type and number of the cations bound. This suggests that the distance and the orientation of chromophores do not change significantly with the changes of cations. The CD band of the monomeric assembly does not seem to differ considerably from that of the dimeric assembly.

This might be attributed to the large separation of G-quartets in TASQ·Cs<sup>+</sup> which minimizes the interactions of bases due to stacking.

**Figure 3.25.** CD spectra of a 0.2 mM solution of TASQ 10c and TASQ 10c·cation assemblies in chloroform.



### 3.4 Summary and Conclusions

In the present Chapter, we demonstrated the affinity of TASQs for cations, and explored the structural consequences of cation-complexation. Four cations (Na<sup>+</sup>, Sr<sup>2+</sup>, K<sup>+</sup>, and Cs<sup>+</sup>) were selected and NMR spectroscopic methods were utilized to investigate the resulting structures. <sup>1</sup>H NMR and <sup>1</sup>H-<sup>1</sup>H NOESY analysis were performed on each assembly and confirmed the presence of G-quartet units. Diffusion NMR experiments were undertaken and the sizes of the complexes were determined. The combined NMR analyses suggest the presence of an isolated

unimolecular G-quartet in the absence of cations or in the presence of Na<sup>+</sup> or Sr<sup>2+</sup>. In the presence of Cs<sup>+</sup>, we found a heterodimer in solution. In the presence of K<sup>+</sup>, both a heterodimer and a unimolecular G-quartet were found. A deeper analysis revealed the role of the anions in the formation and stabilization of these structures. Both dimeric and monomeric assemblies, as noted for biological systems, acted like metal-ion or counter-ion switches. Switching from Cs<sup>+</sup> to Na<sup>+</sup>, for example, induced a conformational change from a dimer to a monomer. Selectivities were evaluated and Sr<sup>2+</sup> was found to be coordinated more strongly than the other cations. Interconversion of structures was also investigated. The sodium-bound complex was converted to cation-free TASQ by addition of cryptand, but the cation-free TASQ remained unaffected under similar conditions. Both experiments were monitored via <sup>1</sup>H NMR. The following Chapter will provide a summary of the studies conducted towards the synthesis of hydrophilic TASQs.

## 3.5 Experimental

### 3.5.1 General

The experimental conditions employed for the NMR spectroscopic measurements are outlined in Chapter 2. An Esquire~LC ion trap mass spectrometer equipped with an electrospray ion source was utilized for the ESI-MS experiments. The solvent system was  $\text{CHCl}_3/\text{CH}_3\text{CN}$  and the sample solution concentration was about 25  $\mu\text{M}$ . It was infused into the ion source by a syringe pump at a flow rate of 10  $\mu\text{l}/\text{min}$ . Mass spectra were acquired in the positive ion mode. MALDI mass spectra were obtained on Bruker Biflex IV MALDI-TOF equipped with nitrogen laser. The samples were dissolved in dichloromethane and DCTB was used as matrix. The solutions of the sample (~1mg/mL) and matrix (20mg/mL) were mixed in the ration of 1:1 to 1:10 and 1 $\mu\text{l}$  of final mixture was deposited onto the sample target. MALDI mass spectra were acquired in the positive reflectron mode with delay extraction. Spectra were normally obtained by averaging 100 laser shots. Calibration was performed externally using various peptides. A Varian SpectrAA-300 atomic absorption spectrometer equipped with a graphite furnace and a GTA autosampler, and a Zeeman effect background correction system was used for sodium determination. A Varian Techtron Na-K hollow cathode lamp was operated at 5 mA with a slit of 0.2 nm. The wavelength was 589.6 nm. Peak heights were used for signal evaluation. All chemicals used for AA analysis were of analytical-reagent grade, unless otherwise specified. De-ionized water was produced by a Milli-Q Academic system (Millipore, Bedford, USA) and the Na standard came as 1000  $\mu\text{g}/\text{ml}$  in 5%  $\text{HNO}_3$ , Specpure from Alfa Aesar.

### 3.5.2 Picrate Extraction Experiments

Picrate salts were synthesized following the literature method.<sup>57</sup> In a typical solid-liquid extraction experiment 0.01 mmol of **10c** was mixed with 0.04 mmol of sodium picrate and stirred in chloroform for 6 days. The resulting mixture was centrifuged and the solution was filtered through a piece of glass wool (packed in a Pasteur pipette). The product was isolated as a yellow solid after evaporation of chloroform at room temperature. Chloroform/aqueous picrate extraction was not practical, as precipitation resulted.

**TASQ 10c·Na<sup>+</sup>**: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 11.55 (s, 4 H, NH), 8.69 (s, 2 H, Picrate), 7.79 (s, 4 H, H<sub>c</sub>), 7.78 (s, 4 H, H<sub>8</sub>), 6.65 (s, 4 H, ArH), 5.88 (d, 4H, H<sub>1'</sub>), 5.67 (d, 4 H, H<sub>b</sub>), 5.19 (dd, 4 H, H<sub>2'</sub>), 5.18 (dd, 4 H, H<sub>3'</sub>), 4.96 (dd, 4 H, H<sub>5'b</sub>), 4.76 (d, 4 H, H<sub>a</sub>), 4.61 (dd, 4 H, H<sub>5'a</sub>), 4.53 (t, 4 H, H<sub>d</sub>), 4.50 (d, 4 H, H<sub>out</sub>), 4.31 (m, 4 H, H<sub>4'</sub>), 3.93 (d, 4 H, H<sub>in</sub>), 2.30 (m, 8 H, CH<sub>2</sub> cavitand feet), 1.48 (s, 12 H, CH<sub>3</sub> sugar), 1.39 (m, 8 H, CH<sub>2</sub> feet), 1.29 (s, 12 H, CH<sub>3</sub> sugar), 1.23-1.30 (m, 64 H, CH<sub>2</sub> feet), 0.84 (t, 12 H, CH<sub>3</sub> feet).

**TASQ 10c·Sr<sup>2+</sup>**: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 11.15 (s, 4 H, NH), 8.70 (s, 2 H, Picrate), 7.87 (s, 4 H, H<sub>c</sub>), 7.80 (s, 4 H, H<sub>8</sub>), 6.62 (s, 4 H, ArH), 5.86 (d, 4H, H<sub>1'</sub>), 5.54 (d, 4 H, H<sub>b</sub>), 5.24 (dd, 4 H, H<sub>3'</sub>), 5.00 (dd, 4 H, H<sub>2'</sub>), 4.96 (dd, 4 H, H<sub>5'b</sub>), 4.78 (d, 4 H, H<sub>a</sub>), 4.63 (dd, 4 H, H<sub>5'a</sub>), 4.54 (t, 4 H, H<sub>d</sub>), 4.46 (d, 4 H, H<sub>out</sub>), 4.30 (m, 4 H, H<sub>4'</sub>), 3.95 (d, 4 H, H<sub>in</sub>), 2.30 (m, 8 H, CH<sub>2</sub> cavitand feet), 1.48 (s, 12 H, CH<sub>3</sub> sugar), 1.39 (m, 8 H, CH<sub>2</sub> feet), 1.29 (s, 12 H, CH<sub>3</sub> sugar), 1.23-1.30 (m, 64 H, CH<sub>2</sub> feet), 0.84 (t, 12 H, CH<sub>3</sub> feet).

**TASQ 10c·Cs<sup>+</sup>**: (partially assigned, region 6-12), <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 11.41 (s, 4 H, NH), 10.83 (s, 4 H, NH\*), 8.87 (s, 4 H, Picrate), 7.70 (s, 4 H, H<sub>c</sub>), 7.42 (s, 4 H, H<sub>c</sub>\*), 7.44 (s, 4 H, H<sub>8</sub>), 7.17 (s, 4 H, H<sub>8</sub>\*), 6.71 (s, 4 H, ArH), 6.51 (s, 4 H, ArH\*)

### 3.5.3 PFG NMR Spectroscopic Experiments

For the general conditions see the section 2.5.7.1.

#### 3.5.3.1 Experimental (Bruker) Parameters for PFG NMR Experiments

The following mixtures were prepared and subjected to analysis: **16c-10c·Na<sup>+</sup>**, **16c-10c·Sr<sup>2+</sup>**, and **10c-10c·Cs<sup>+</sup>** in CDCl<sub>3</sub>. The sample and standard were kept in the same NMR tube and the experimental conditions were the same for all of the assemblies. No changes were observed in the <sup>1</sup>H NMR spectra upon long term storage.

NS [scan]: 32	TD (F2): 8K	TD (F1): 16
D21 [delay]: 5ms	D20 [Δ]: 100 ms	DS: 8
P30 [δ/2]: 1.5 ms	P19: 0.75 ms	GPZ6 [G]: 100%
GPZ7: -17.13%	GPZ8: -13.17	

### 3.5.4 Atomic Absorption Experiments

The sodium content of the cation-free TASQ **10c** and TASQ **10c**·Na<sup>+</sup> was determined by atomic absorption spectroscopy (see the general section for the reagents and instrumentation). 1 mg of each sample was digested in concentrated nitric acid. The solution was evaporated and the residue was dissolved in 2 mL of nitric acid 1% v/v to make a stock solution. For the cation-free TASQ, 2 μL of the stock solution was diluted 6-fold with nitric acid 1% v/v and injected into the furnace (total volume = 12 μL). For TASQ **10c**·Na<sup>+</sup>, the original stock solution had to be diluted again 200 fold due to the high concentration of sodium. Then, 2 μL of the stock solution was mixed with 10 μL of the blank (nitric acid 1% v/v) and injected into the furnace (total volume = 12 μL). The results are summarized in Table 3.5. Each number is the average of two measurements. The results of AA analysis are in a very good agreement with the results of NMR analysis and confirm the presence of TASQ **10c** as a different entity from TASQ **10c**·Na.

A slightly excess amount of sodium detected for TASQ **10c**·Na (Na content = 1.13eq.) is within the experimental error ( $\pm 10\%$ ).

**Table 3.5.** Na content ( $\mu\text{g/g}$ ) of TASQ samples.

Sample	Na <sup>+</sup> content (ppm)	RSD (n=2)
TASQ <b>10c</b>	250	5.4
TASQ <b>10c</b> ·Na <sup>+</sup>	9456	6.3

### 3.6 References

- (1) Neidle, S.; Balasubramanian, S. *Quadruplex nucleic acids*; RSC Pub.: Cambridge, 2006.
- (2) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.
- (3) Keniry, M. A. *Biopolymers* **2000**, *56*, 123-146.
- (4) Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491-511.
- (5) Sen, D.; Gilbert, W. *Methods Enzymol.* **1992**, *211*, 191-199.
- (6) Włodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. *J. Phys. Chem. B* **2005**, *109*, 3594-3605.
- (7) Wong, A.; Wu, G. *J. Am. Chem. Soc.* **2003**, *125*, 13895-13905.
- (8) van Leeuwen, F. W. B.; Davis, J. T.; Verboom, W.; Reinhoudt, D. N. *Inorg. Chim. Acta* **2006**, *359*, 1779-1785.
- (9) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. *Science* **1994**, *265*, 520-524.
- (10) Wong, A.; Ida, R.; Wu, G. *Biochem. Biophys. Res. Commun.* **2005**, *337*, 363-366.
- (11) Wong, A.; Fettinger, J. C.; Forman, S. L.; Davis, J. T.; Wu, G. *J. Am. Chem. Soc.* **2002**, *124*, 742-743.
- (12) Deng, H.; Braunlin, W. H. *J. Mol. Biol.* **1996**, *255*, 476-483.
- (13) Ma, L.; Iezzi, M.; Kaucher, M. S.; Lam, Y. F.; Davis, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 15269-15277.
- (14) Sket, P.; Crnugelj, M.; Plavec, J. *Nucleic Acids Res.* **2005**, *33*, 3691-3697.
- (15) Sen, D.; Gilbert, W. *Nature* **1990**, *344*, 410-414.
- (16) Hardin, C. C.; Henderson, E.; Watson, T.; Prosser, J. K. *Biochemistry* **1991**, *30*, 4460-4472.
- (17) McPhee, M. M.; Kern, J. T.; Hoster, B. C.; Kerwin, S. M. *Bioorg. Chem.* **2000**, *28*, 98-118.
- (18) Roomans, G. M.; vonEuler, A. *Cell. Biol. Int.* **1996**, *20*, 103-109.
- (19) Kaucher, M. S.; Davis, J. T. *Tetrahedron Lett.* **2006**, *47*, 6381-6384.

- (20) Shannon, R. D.; Prewitt, C. T. *Acta Crystallogr., Sect. B: Struct. Sci.* **1969**, *B* 25, 925.
- (21) Pinnavaia, T. J.; Marshall, C. L.; Mettler, C. M.; Fisk, C. I.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1978**, *100*, 3625-3627.
- (22) Forman, S. L.; Fettingner, J. C.; Pieraccini, S.; Gottareli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067.
- (23) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171-182.
- (24) Marlow, A. L.; Mezzina, E.; Spada, G. P.; Masiero, S.; Davis, J. T.; Gottarelli, G. *J. Org. Chem.* **1999**, *64*, 5116-5123.
- (25) Aboulela, F.; Murchie, A. I. H.; Lilley, D. M. J. *Nature* **1992**, *360*, 280-282.
- (26) Wang, Y.; Patel, D. J. *Biochemistry* **1992**, *31*, 8112-8119.
- (27) Liu, X. Y.; Kwan, I. C. M.; Wang, S. N.; Wu, G. *Org. Lett.* **2006**, *8*, 3685-3688.
- (28) Nikan, M.; Sherman, J. C. *Angew. Chem., Int. Ed.* **2008**, *47*, 4900-4902.
- (29) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A. *Angew. Chem., Int. Ed.* **2000**, *39*, 1300-1303.
- (30) Gottarelli, G.; Masiero, S.; Spada, G. P. *J. Chem. Soc., Chem. Commun.* **1995**, 2555-2557.
- (31) Mezzina, E.; Mariani, P.; Itri, R.; Masiero, S.; Pieraccini, S.; Spada, G. P.; Spinuzzi, F.; Davis, J. T.; Gottarelli, G. *Chem.-Eur. J.* **2001**, *7*, 388-395.
- (32) Aggerholm, T.; Nanita, S. C.; Koch, K. J.; Cooks, R. G. *J. Mass Spectrom.* **2003**, *38*, 87-97.
- (33) Shi, X.; Fettingner, J. C.; Davis, J. T. *Angew. Chem., Int. Ed.* **2001**, *40*, 2827-2831.
- (34) Chen, F. M. *Biochemistry* **1992**, *31*, 3769-3776.
- (35) Kankia, B. I.; Marky, L. A. *J. Am. Chem. Soc.* **2001**, *123*, 10799-10804.
- (36) Mao, X. A.; Marky, L. A.; Gmeiner, W. H. *J. Biomol. Struct. Dyn.* **2004**, *22*, 25-33.
- (37) Deng, J. P.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13665-13670.
- (38) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189-200.

- (39) Patel, D. J.; Kozlowski, S. A.; Nordheim, A.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 1413-1417.
- (40) Feigon, J.; Wang, A. H. J.; Vandermarel, G. A.; Vanboom, J. H.; Rich, A. *Nucleic Acids Res.* **1984**, *12*, 1243-1263.
- (41) Talanova, G. G.; Elkarim, N. S. A.; Talanov, V. S.; Hanes, R. E.; Hwang, H. S.; Bartsch, R. A.; Rogers, R. D. *J. Am. Chem. Soc.* **1999**, *121*, 11281-11290.
- (42) Olsher, U.; Feinberg, H.; Frolow, F.; Shoham, G. *Pure Appl. Chem.* **1996**, *68*, 1195-1199.
- (43) Inoue, Y.; Fujiwara, C.; Wada, K.; Tai, A.; Hakushi, T. *J. Chem. Soc., Chem. Commun.* **1987**, 393-394.
- (44) Bohmer, V.; Dalla Cort, A.; Mandolini, L. *J. Org. Chem.* **2001**, *66*, 1900-1902.
- (45) Kaucher, M. S.; Lam, Y. F.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *Chem.-Eur. J.* **2004**, *11*, 164-173.
- (46) Ross, W. S.; Hardin, C. C. *J. Am. Chem. Soc.* **1994**, *116*, 6070-6080.
- (47) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. *Biochemistry* **1996**, *35*, 15383-15390.
- (48) Chitry, F.; Pellet-Rostaing, S.; Nicod, L.; Gass, J. L.; Foos, J.; Guy, A.; Lemaire, M. *J. Phys. Chem. A* **2000**, *104*, 4121-4128.
- (49) DeFilippi, I.; Yates, S.; Sedath, R.; Straszewski, M.; Andren, M.; Gaita, R. *Sep. Sci. Technol.* **1997**, *32*, 93-113.
- (50) Hobbs, D. T.; Barnes, M. J.; Pulmano, R. L.; Marshall, K. M.; Edwards, T. B.; Bronikowski, M. G.; Fink, S. D. *Sep. Sci. Technol.* **2005**, *40*, 3093-3111.
- (51) Lehn, J. M.; Sauvage, J. P. *J. Am. Chem. Soc.* **1975**, *97*, 6700-6707.
- (52) Ghossoub, A.; Lehn, J. M. *Chem. Commun. (Cambridge, U. K.)* **2005**, 5763-5765.
- (53) Pieraccini, S.; Masiero, S.; Pandoli, O.; Samori, P.; Spada, G. P. *Org. Lett.* **2006**, *8*, 3125-3128.
- (54) Gottarelli, G.; Masiero, S.; Spada, G. P. *Enantiomer* **1998**, *3*, 429-438.
- (55) Gray, D. M.; Wen, J. D.; Gray, C. W.; Repges, R.; Repges, C.; Raabe, G.; Fleischhauer, J. *Chirality* **2008**, *20*, 431-440.
- (56) Novy, J.; Bohm, S.; Kralova, J.; Kral, V.; Urbanova, M. *Biopolymers* **2008**, *89*, 144-152.

(57) Wong, K. H.; Ng, H. L. *J. Coord. Chem.* **1981**, *11*, 49-55.

## **CHAPTER 4: Towards Hydrophilic TASQs\***

### **4.1 Introduction**

Chapter 2 discussed the design, synthesis, and solution behavior of lipophilic TASQs. Chapter 3 described the preparation and characterization of lipophilic TASQ-cation assemblies. This Chapter describes efforts toward the synthesis of a hydrophilic TASQ. The synthesis of starting materials was completed, but time constraints precluded completion of the syntheses as work described in Chapters 2 and 3 became the main focus of the thesis work. Suggestions for future work are made at the end of this Chapter.

### **4.2 Rationale for the Study of Hydrophilic TASQs**

Biological reactions take place in aqueous environment. G-quadruplex structures are present in biologically important regions<sup>1-4</sup> and have been linked to many biological and pathological processes such as cancer,<sup>5</sup> aging<sup>6,7</sup> and genetic diseases.<sup>8</sup> A natural extension of the studies carried out in lipophilic systems is the hydrophilic TASQ project which is aimed to explore the behavior of the TASQs in aqueous systems. Self-assembled architectures which are held together by non-covalent interactions are solvent dependent, since their structure and stability are generally affected by a change of solvent.

---

\* "A version of this Chapter will be submitted for publication. Nikan, M.; Sherman, J. C., Synthesis of Phosphate-Footed (*O*-Tetrapropargyl) Cavitand."

The first challenge before studying the behavior of the hydrophilic TASQs is their synthesis. Our initial goal is to synthesize an all-parallel G-quadruplex in which all nucleotides are in 5'-3' orientations with respect to the cavitand. According to Piccialli et al., most of the biologically important systems are of the antiparallel type.<sup>9</sup> However, some important parallel-stranded systems such as aptamers with anti-HIV activity<sup>10,11</sup> and hematoporphyrin binding aptamers<sup>12</sup> have been identified. There are also proteins and ligands which interact with or stabilize the parallel systems.<sup>9,13</sup> It has been reported that *in vitro* formation of parallel G-quadruplexes is slow.<sup>14</sup> Our proposed system facilitates the formation of G-quadruplexes at low concentrations and might pave the way for their biological applications.

### **4.3 Results and Discussion**

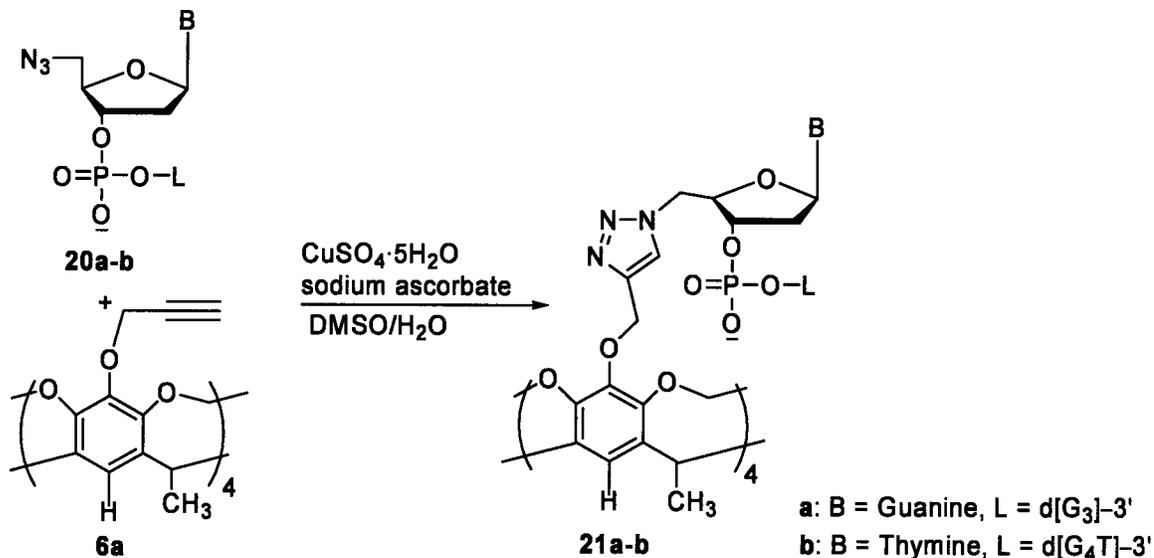
Two approaches were taken to incorporate water solubility into TASQs: One approach uses the water-solubilizing phosphate groups of oligonucleotides (section 4.3.1), while the second approach uses phosphate-footed cavitands (section 4.3.2). Regarding methods to link the guanines to the cavitands, both “click” chemistry<sup>15</sup> (sections 4.3.1 and 4.3.2) and phosphoramidite methods<sup>16</sup> (section 4.3.3) were explored.

#### **4.3.1 The Oligonucleotide Approach**

This approach involves the coupling of an azide-functionalized oligonucleotide to a tetrapropargyl ether cavitand (Scheme 4.1). The cavitand template does not need to be water-

soluble since the solubility of the oligonucleotides (12 or 16 charges per unit TASQ) is expected to be sufficient to drag the whole system ( $\sim$ MW = 2086 or 2696) into water.

**Scheme 4.1.** Proposed synthesis of an all parallel hydrophilic TASQ with the sequences of 5'-d[TG<sub>4</sub>T]-3' and 5'-d[G<sub>4</sub>]-3' using the oligonucleotide approach.



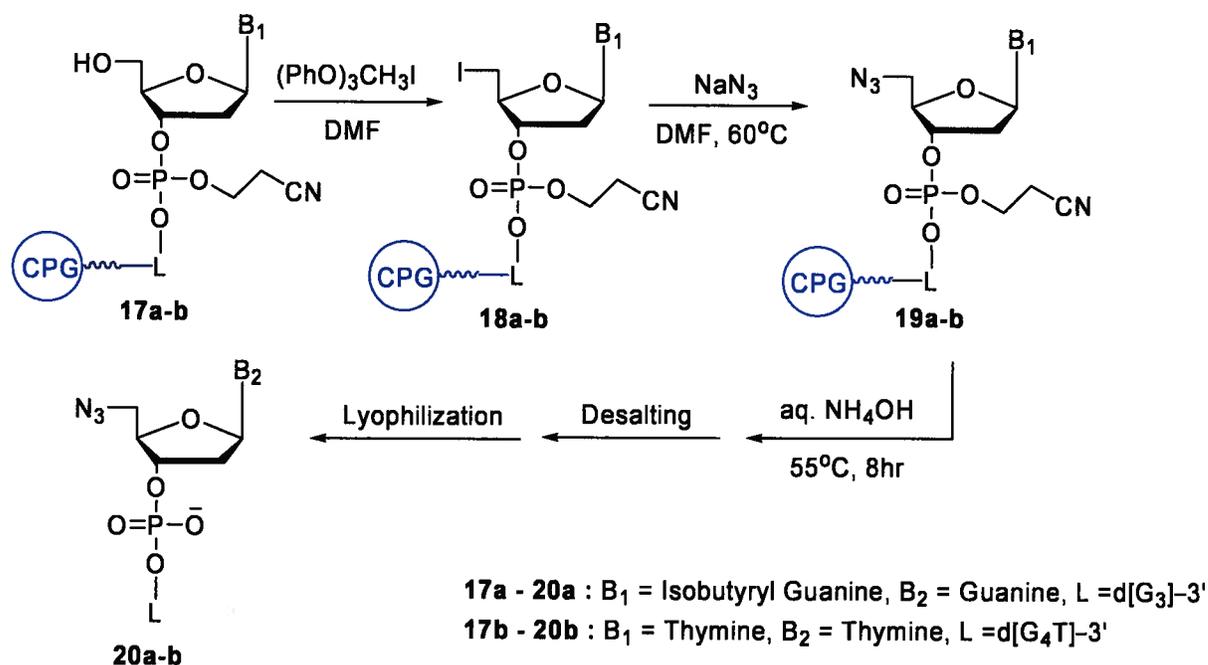
Two guanine-rich sequences, 5'-d[N<sub>3</sub>-G<sub>4</sub>]-3' (**20a**) and 5'-d[N<sub>3</sub>-TG<sub>4</sub>T]-3' (**20b**) were selected for this study. The thymine bases of **20b** were included to enhance the solubility and to prevent the inter-TASQ stacking of the G-quadruplexes. Such stacking is common in the quadruplexes containing a terminal G-quartet and might lead to the formation of higher order structures.<sup>17-19</sup> The incorporation of azide functionalities onto the 5' termini of the oligonucleotides and the subsequent “click” reaction was expected to provide a linker of the same length as the ones on the lipophilic TASQs. Our group’s studies on TASPes (template-assembled synthetic peptides) have shown that the length of the linker has an important impact on the structure and properties of the TASPes.<sup>20,21</sup> Short linkers were found to induce a strain in

the H-bonding network of the TASP s while long flexible linkers decreased the stabilizing effect of the cavitand. Therefore, we maintained the triazole linker that had worked well in the previous studies. It should also be noted that the azide group is not compatible with the phosphoramidite group and cannot be introduced during the solid-phase synthesis.<sup>22,23</sup> Thus, it was obtained via a post-synthesis modification (section 4.3.1.1). The coupling reaction shown in Scheme 4.1 will be discussed in section 4.3.1.2, but in short, success has not yet been achieved in the synthesis of compounds **21a-b**.

#### 4.3.1.1 Solid-Phase Functionalization of the Oligonucleotides

Miller and Kool recently introduced a general method for the 5'-functionalization of oligonucleotides.<sup>24,25</sup> The synthesis can be performed semi-automatically or manually. We took the manual approach (the Chemistry Department of the UBC is not equipped with a DNA synthesizer). Following this method, **17a-b** (purchased from the University of Calgary) which were attached to the CPG (controlled-pore glass) inside a DNA synthesizing column were first treated with a solution of methyltriphenoxyposphonium iodide. This incorporates an electrophilic site (-CH<sub>2</sub>I) onto the 5' position. The iodinated products were then reacted with a saturated solution of sodium azide in DMF. Subsequent cleavage of the oligonucleotides from the solid support, deprotection, butanol precipitation and purification gave **20a-b**. The synthetic steps are illustrated in Scheme 4.2.

**Scheme 4.2.** The solid-phase functionalization of **17a-b**.



#### 4.3.1.2 The Coupling Reaction of 5'-Azido Oligonucleotides **20a-b**

The final step in the synthesis of hydrophilic TASQs **21a-b** is the coupling of the oligonucleotides **20a-b** to the cavitand template **6a**. A variety of conditions have been tested so far on one of the sequences (**20a**) with no positive result (Entries 1-10, Table 4.1). Various parameters such as catalyst, molar ratio of the reactants, temperature and steric effects may influence the final “click” reaction. Some, but not all of these parameters have been tested on **20a** and will be discussed here. In each of the following paragraphs we discuss our results in the context of related work in the literature, and make suggestions for future work.

**Table 4.1** Reaction conditions examined for the coupling<sup>a</sup> of **20a** to **6a**.

Entry	<b>6a</b> (mol)	<b>20a</b> (mol)	Cu(II) (mol)	Reducing Agent (mol)	TBTA (mol)	T (°C)	Time (h)	Solvent
1	$2.1 \times 10^{-7}$	<b>20a-CTA</b> $9.6 \times 10^{-7}$	$2.1 \times 10^{-8}$	sodium ascorbate $2.1 \times 10^{-7}$	$2.1 \times 10^{-8}$	70	20	DMSO
2	$6.5 \times 10^{-9}$	$3.1 \times 10^{-8}$	$6.5 \times 10^{-10}$	sodium ascorbate $6.5 \times 10^{-9}$	–	60	48	H <sub>2</sub> O:THF 1:1
3	$6.5 \times 10^{-9}$	$3.1 \times 10^{-8}$	$6.5 \times 10^{-10}$	sodium ascorbate $6.5 \times 10^{-9}$	–	70	20	DMSO:H <sub>2</sub> O 1:1
4	$3.2 \times 10^{-8}$	$1.5 \times 10^{-7}$	$3.2 \times 10^{-7}$	sodium ascorbate $3.2 \times 10^{-6}$	$3.2 \times 10^{-7}$	90	48	DMSO:H <sub>2</sub> O:t-BuOH 3:3:1
5	$3.2 \times 10^{-8}$	$1.5 \times 10^{-7}$	$3.2 \times 10^{-9}$	TCEP $3.2 \times 10^{-8}$	$3.2 \times 10^{-9}$	60	24	DMSO:H <sub>2</sub> O:t-BuOH 3:3:1
6	$3.2 \times 10^{-8}$	$1.5 \times 10^{-7}$	$3.2 \times 10^{-9}$	TCEP $3.2 \times 10^{-8}$	$3.2 \times 10^{-9}$	25	48	DMSO:H <sub>2</sub> O:t-BuOH 3:3:1
7	$2.1 \times 10^{-7}$	$9.6 \times 10^{-7}$	$2.1 \times 10^{-8}$	Cu(0) excess	–	70	20	DMSO:H <sub>2</sub> O 1:1
8	$2.1 \times 10^{-7}$	$9.6 \times 10^{-7}$	Cu(I) $2.1 \times 10^{-8}$	sodium ascorbate $2.1 \times 10^{-7}$	$2.1 \times 10^{-8}$	70	24	DMSO:H <sub>2</sub> O 1:1
9	$6.5 \times 10^{-9}$	$3.1 \times 10^{-8}$	$6.5 \times 10^{-10}$	sodium ascorbate $6.5 \times 10^{-9}$	$6.5 \times 10^{-10}$	25	24	DMSO:H <sub>2</sub> O 1:1
10	$3.2 \times 10^{-8}$	$1.5 \times 10^{-7}$	$3.2 \times 10^{-9}$	sodium ascorbate $3.2 \times 10^{-8}$	$3.2 \times 10^{-9}$	90	48	DMSO:H <sub>2</sub> O:t-BuOH 3:3:1

<sup>a</sup> Due to the small amount, the reaction mixtures were only analyzed by MALDI-MS. It turned out, however, that finding a matrix that works well for both the cavitaand and oligonucleotide is challenging. For example, 3-HPA, a matrix widely used for oligonucleotides transmits less energy to the cavitaand, while DHB, a matrix suitable for the detection of cavitaands, induces intense fragmentation of the oligonucleotide. For the current set of experiments, the crude reaction mixtures obtained from Entries 1-10 were mixed with 3-HPA:imidazole matrix:co-matrix combination and analyzed.

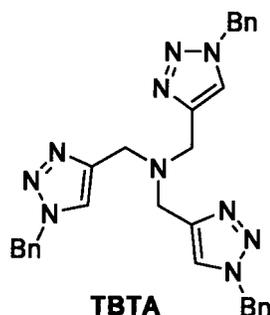
The first parameter examined was solubility. We first attempted to use the reaction conditions of the lipophilic TASQs (section 2.5.4). Since **20a** is almost insoluble in DMSO, it was converted to its organic soluble CTA salts. Cetyltrimethylammonium bromide (CTAB) is a cationic detergent which is used to solubilize oligonucleotides in organic phases.<sup>26</sup> However, the coupling reaction was not successful (Entry 1, Table 4.1). Besides this attempt, the rest of the reactions were carried out in water/co-solvent (DMSO/*t*-BuOH or THF) mixtures which provided a more or less homogeneous system. In an early paper published by Sharpless et al., the authors stated that the reactants do not have to be completely soluble and the reaction should proceed well as long as good stirring is maintained.<sup>15</sup>

We have also examined the role of the catalyst. In Chapter 1 we noted that a ratio of 0.02 eq. of Cu(II) per 1 eq. of azide is sufficient to obtain the product. However, this did not give product in the current study (Entry 2-3, Table 4.1). Reports concerning the “click” reaction of oligonucleotides vary considerably regarding the optimum ratio of catalyst to reagents. A recent report suggests using a large excess of Cu(II) (10-25 eq. per 1 eq. of oligonucleotide).<sup>23</sup> Such ratios are significantly higher than those attempted for this project. Perhaps such a large excess of catalyst is needed, but this may be also problematic for the following reason. The Cu(I) catalyst which is required for the “click” reaction is unstable in water and readily disproportionates to Cu(0) and Cu(II). Thus, it is typically generated *in situ* from Cu(II) and sodium ascorbate. One of the by-products of the latter catalytic pair is a hydroxyl radical, which can potentially damage the DNA (Haber-Weiss reaction).<sup>27-30</sup> Replacing sodium ascorbate with other reducing agents such as tris(2-carboxyethyl)phosphine hydrochloride (TCEP)<sup>31</sup> (Entry 5-6, Table 4.1) and Cu(0) (Entry 7, Table 4.1) or direct introduction of Cu(I)<sup>15</sup> (Entry 8, Table 4.1) or changing the temperature (Entry 9-10, Table 4.1) gave no product. Recently, a number of

copper-based catalytic systems such as copper-in-charcoal (Cu/C)<sup>32</sup> or copper(I) zeolites<sup>33</sup> have been introduced which have not yet been examined by us.

The use of the TBTA ligand (Figure 4.1, commercially available) which is believed to accelerate the reaction<sup>31,34</sup> has also been reported by Alvira and Eritja to have little effect on the yield.<sup>23</sup> The TBTA ligand is used to stabilize the Cu(I) and also to prevent the oxidative coupling of the terminal alkynes. In the preliminary reactions conducted on **6a** and **20a**, we observed no product upon addition of TBTA at the attempted concentration to the reaction mixture (Entry 1, 4-6, 8-10, Table 4.1).

**Figure 4.1.** Tris-(benzyltriazolylmethyl)amine (TBTA)



Surprisingly, in a model reaction, we found that **20a** undergoes a “click” reaction with small alkynes with no apparent degradation.<sup>†</sup> However, reaction of the cavitand was not observed. Recall that cavitand **6a** also undergoes the “click” reaction with other nucleotides (section 2.5.4). So, why does the “click” reaction not function for the formation of the hydrophilic TASQs? Our observations, analyzed in the light of the available literature, suggest that the geometry of the cavitand might also play a role. Recently, Ryu and Zhao reported an

<sup>†</sup> The reaction of **20a** with 2-Methyl-3-butyn-2-ol afforded the corresponding “click” product which was characterized by ESI-MS.

unsuccessful attempt at a “click” reaction of a tetrapropargyl ether calixarene to a water-soluble group.<sup>35</sup> Calixarenes are structurally similar to the cavitands but are flexible. Although the authors attributed the poor results to the large solubility difference between the reactants, others disagree. Van Maarseveen and colleagues have suggested that the proximity of the alkyne groups in calixarene and the preference of the catalyst for the alkynes over azides are the reasons why this reaction fails.<sup>36</sup> The Cu(I) atom supposedly undergoes a reaction with the first alkyne group to form a copper acetylide species needed for the “click” reaction. However, this also places the catalyst in the vicinity of the other neighboring alkynes which can saturate the catalyst by further complexation. As a result of this process, the azide group doesn’t bind to the Cu(I) atom and the reaction fails, according to van Maarseveen. Considering the geometry of the cavitand, this might also be an explanation of why the coupling reactions did not occur. However, a detailed mechanistic investigation is needed to confirm this hypothesis and to explain why it has only been observed for the coupling of the cavitand to oligonucleotides.<sup>‡</sup> The “click” reaction is generally considered a robust reaction that works well under a variety of conditions. Reports regarding problematic “click” reactions are rare.<sup>36</sup> Thus, aside from the problem of synthesizing the products it was interesting for us to follow this line of research and to investigate the scope of this reaction.

In conclusion, we have provided an overview of the results obtained so far from the “click” reaction of the cavitand to the oligonucleotides. One or a combination of factors might be responsible for the poor results. If the latter hypothesis is confirmed - that the saturation of the catalyst is the reason for the failure of the reactions - then switching the alkyne and azide functionalities on the cavitand and oligonucleotide should be sufficient to overcome the problem. If not, there are a number of conditions which can be explored using the current set of reactants.

---

<sup>‡</sup> Reaction of **20a** with monopropargyl ether cavitand might be a quick way to test this hypothesis.

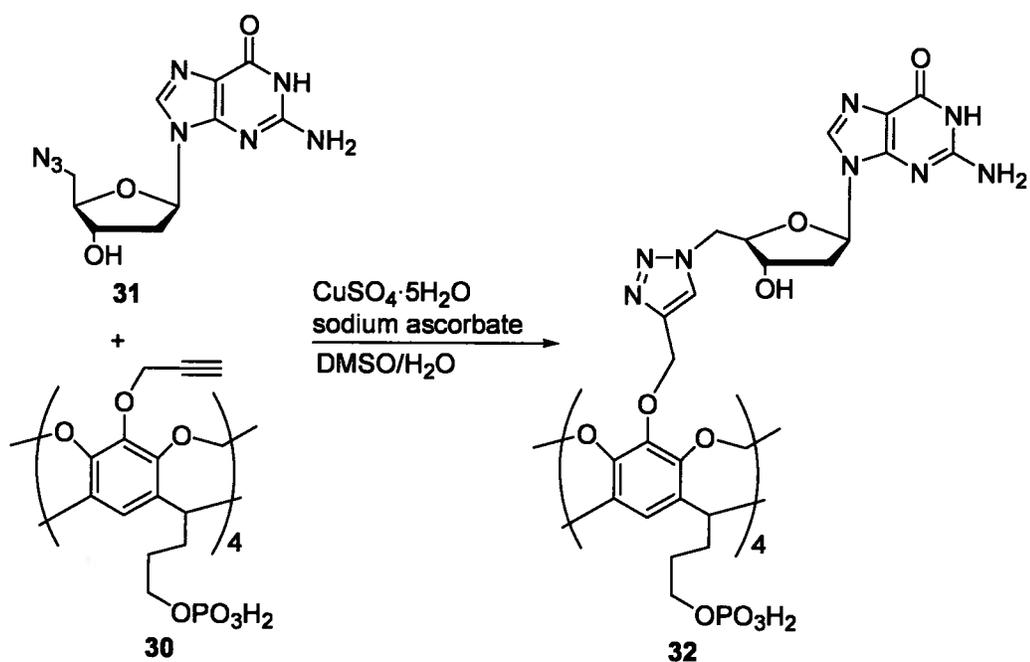
The priority should be given to the study of catalyst loading which should be screened in a wider range of concentrations. The utility of compound **20b** needs to be evaluated. This strand is expected to be more soluble and sterically less-hindered due to the presence of thymine bases, and therefore holds some promise. In addition, the possible interference of sterics needs to be investigated using longer linkers. The coupling reaction could be carried out on solid-support. Several other metal-based catalytic systems such as Ru(II),<sup>37</sup> Pt(II) or Pd(II)<sup>38</sup> have been reported which could be examined as alternatives to the copper catalyst. The reaction could also be conducted under microwave radiation to decrease the reaction time. Moreover, it might be possible to analyze the reaction mixtures using chromatographic or electrophoretic methods. These suggestions and future direction will be discussed in Chapter 5.

#### **4.3.2 The Water-Soluble Cavitand Approach**

This approach involves a “click” reaction between a water-soluble phosphate-footed tetrapropargyl ether cavitand and an azide-functionalized guanosine compound (Scheme 4.3). The water-soluble cavitand **30** was synthesized in 9 steps with the solubilizing phosphate groups successfully incorporated onto the lower rims. In addition to the solubility contributions, the phosphate groups are also meant to prevent the interaction of other cavitands from the lower rims. These tail-to-tail interactions, which would be driven by hydrophobic forces (in the absence of the phosphate groups), might lead to the aggregation of the TASQs in aqueous system. The synthesis of the water-soluble cavitand is described in section 4.3.2.1. The final coupling reaction shown below was attempted twice following the condition described in section 2.5.4. The NMR analysis of the reaction mixture was promising and a new set of signals was

observed, but the product was not detected by MALDI-MS. Unfortunately, the coupling experiments had to be stopped at this stage due to insufficient supply of compound **30** and time/effort going to the work described in Chapters 2 and 3. The completion of this synthesis will answer some of the questions posed in section 4.3.1.2 and will clarify whether there is any difference in reactivity of cavitands towards the small (**31**) and large (**20a**) azides under the reaction conditions.

**Scheme 4.3.** Proposed synthesis of a hydrophilic TASQ using the water-soluble cavitand approach.

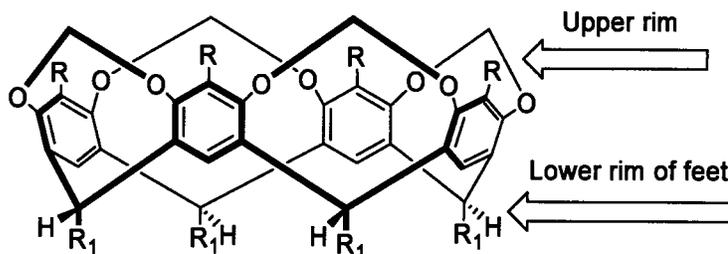


#### 4.3.2.1 Synthesis of Phosphate-Footed Tetrakis (*O*-Propargyl) Cavitand **30**

Selective functionalization of the upper and lower rims of the cavitand (Figure 4.2) is an important aspect of cavitand chemistry. With an appropriate combination of the functional and solubilizing groups, one can have both the advantages of reactivity and solubility. The number

and geometry of the substituents can also be altered to improve the properties of the cavitands for a particular purpose. Adam Mezo, one of the former members of the Sherman lab, developed a synthetic pathway to benzylthiol and benzylbromo cavitands containing pendent phosphate groups in the lower rims.<sup>39</sup> The introduction of the phosphate groups into the cavitand scaffold was attractive to us for the following reasons. First, the incorporation of the phosphate groups into the lower rim of the cavitand would leave the upper rim free for functionalization with the propargyl groups needed for the “click” reaction. Second, the phosphate groups are one of the major constituents of DNA. This would allow us to study the target molecules under the same conditions which are generally used for oligonucleotides without worrying about the complications of the pH and charge (note that the cavitands can also be solubilized with the help of cationic groups<sup>40,41</sup>).

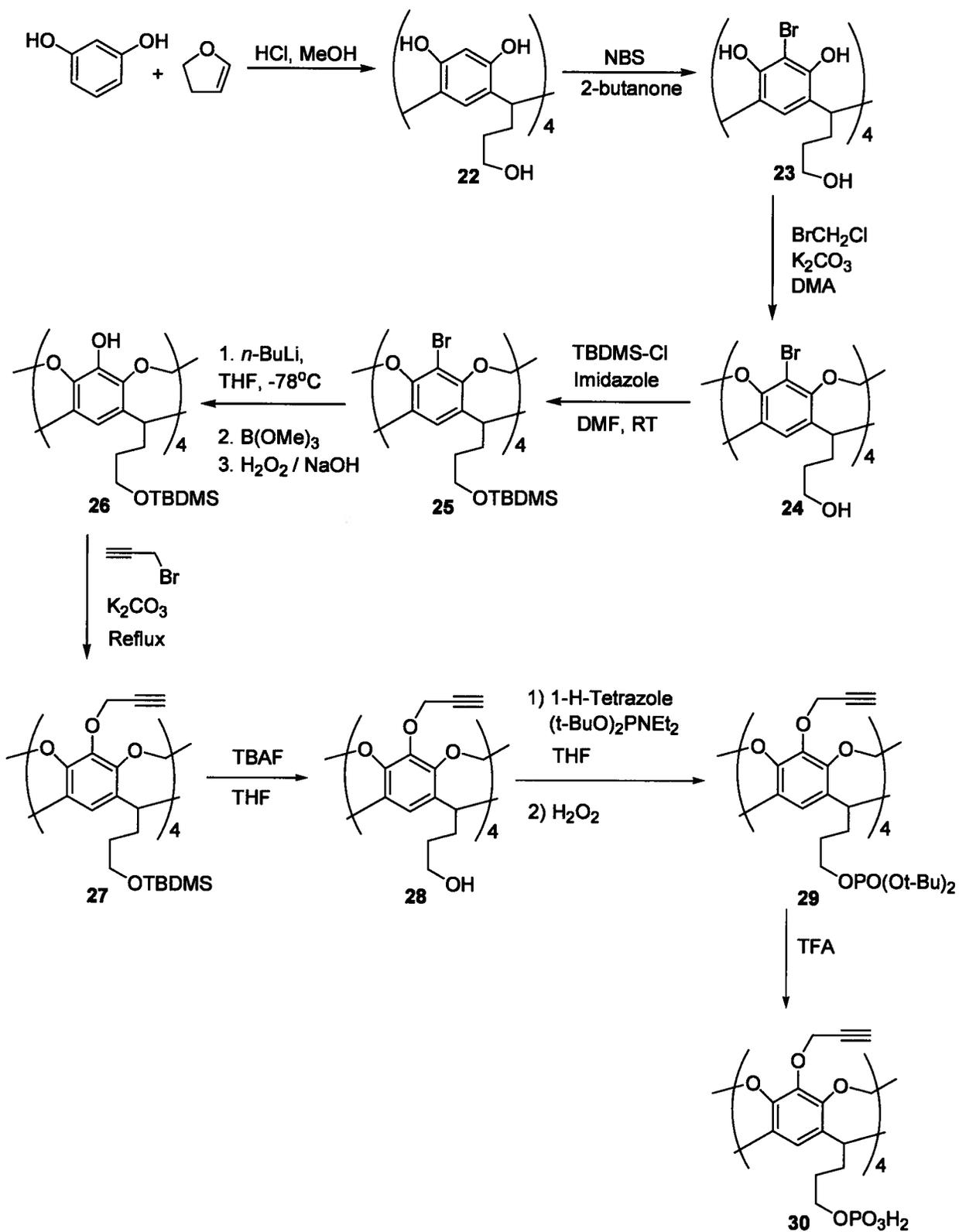
**Figure 4.2.** The upper and lower rims of the cavitand.



The synthesis of compound **30** is illustrated in Scheme 4.4. Compounds **22-24** were synthesized according to a previously reported procedure.<sup>42</sup> Compounds **25-30** have not been previously reported in the literature.<sup>§</sup> The details of the syntheses are described below.

<sup>§</sup> An *O*-benzyl-protected derivative of compound **26** has been synthesized independently by Ayub Jasat, a former member of the Sherman lab. However, neither his results nor the details of his work are available. A sample of this compound was deprotected and utilized for an early synthesis and Ayub’s contribution is greatly appreciated.

**Scheme 4.4.** The synthesis of phosphate-footed tetrakis (*O*-propargyl) cavitand **30**.



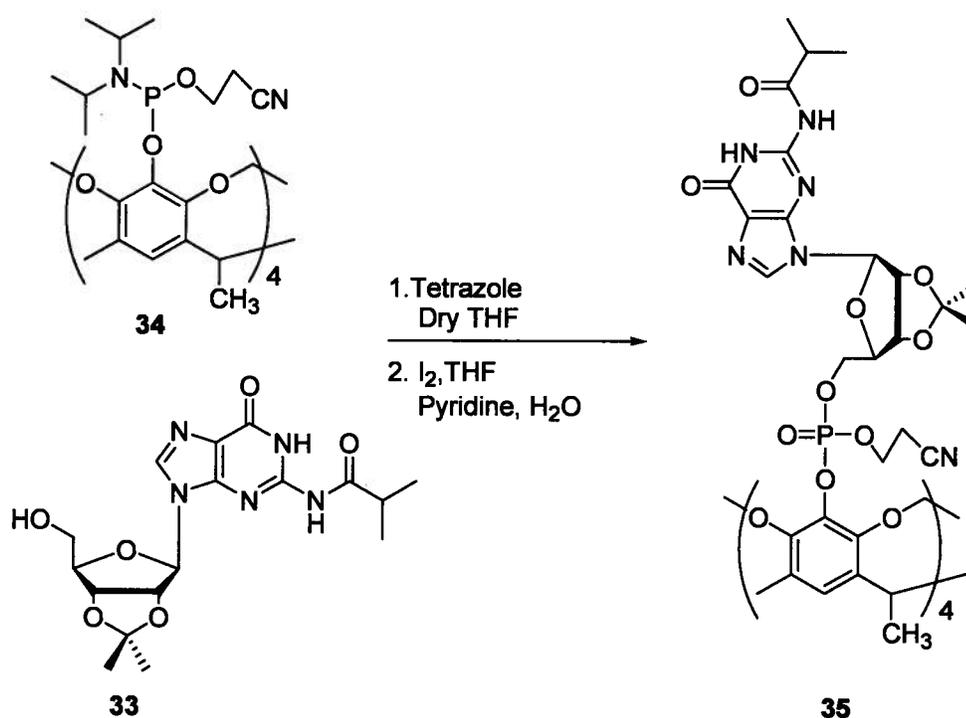
The condensation of resorcinol with 2,3-dihydrofuran under acidic condition gave dodecol **22** after one week in 62% yield. The bromination of **22** with *N*-bromosuccinimide (NBS) afforded tetrabromo dodecol **23** in 80% yield. Compound **23** was then treated with bromochloromethane and potassium carbonate in *N,N*-dimethylacetamide (DMA) to yield tetrabromocavitand **24** in 53% yield. The propargylation of the upper rim of the cavitand was carried out prior to the phosphorylation of the lower rim. Therefore, the hydroxyl groups of the pendant groups had to be protected to avoid interference with the reactions. The protecting group needed to be base-resistant and removable without affecting the propargyl groups. Thus, cavitand **24** was first reacted with tert-butyldimethylsilyl chloride (TBDMS-Cl) to yield the TBDMS-protected cavitand **25** in 67%. Compound **25** was then metalated with *n*-butyllithium, quenched with trimethyl borate and oxidized with NaOH-H<sub>2</sub>O<sub>2</sub> to give TBDMS-protected tetrol **26** (31%). Subsequent reaction of **26** with potassium carbonate and propargyl bromide afforded TBDMS-protected tetrakis (*O*-propargyl) cavitand **27** in 72% yield. Desilylation of **27** with tetrabutylammonium fluoride (TBAF) in dry THF gave compound **28** in 60% yield. The phosphorylation of the lower rim of the cavitand was accomplished in two steps. The tetrazole-catalyzed reaction of **28** with di-tert-butyl *N,N*-diethylphosphoramidite followed by the oxidation with H<sub>2</sub>O<sub>2</sub> gave **29** in 61% yield. Subsequent removal of the tert-butyl groups with TFA afforded the final product **30** in 97% yield.

### 4.3.3 The Phosphoramidite Approach

The solid-phase synthesis of oligonucleotides is carried out in a step-wise manner by linking the protected nucleotides via phosphodiester bonds.<sup>16</sup> Recently, phosphoramidite

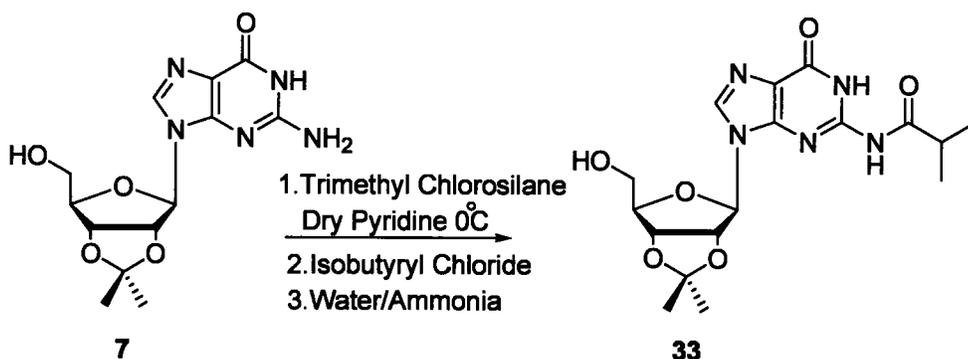
chemistry has been used to prepare nucleotide-calixarene conjugates,<sup>43-45</sup> metal-DNA hybrids,<sup>46-48</sup> and bunched oligonucleotides.<sup>9,49</sup> The following synthesis was intended to use a similar strategy to form a phosphodiester linkage between an *N*-protected guanosine and a tetrakisphosphoramidite cavitand in solution (Scheme 4.5).

**Scheme 4.5.** Proposed synthesis of a hydrophilic TASQ using the phosphoramidite approach.



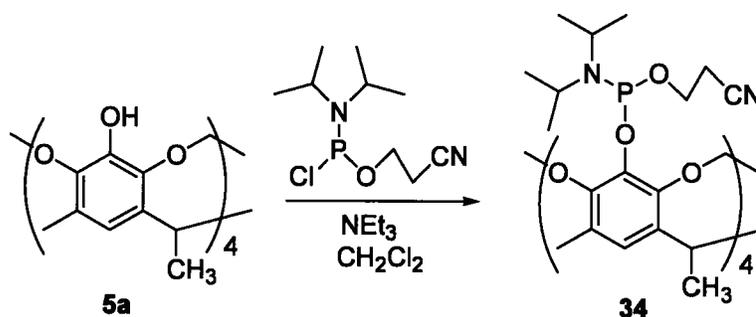
The *N*-protected guanosine **33** was readily synthesized in two steps following Jones's transient protection strategy (Scheme 4.6).<sup>50</sup> The 5'-hydroxyl of 2',3'-*O*-isopropylidene guanosine **7** was first protected as trimethylsilyl ether. The reaction mixture was then treated with isobutyryl chloride to protect the exocyclic amino (NH<sub>2</sub>) group. The silyl protecting groups were then cleaved with an aqueous ammonium hydroxide solution to give **33** in 71% yield.

**Scheme 4.6.** The transient protection of 7.



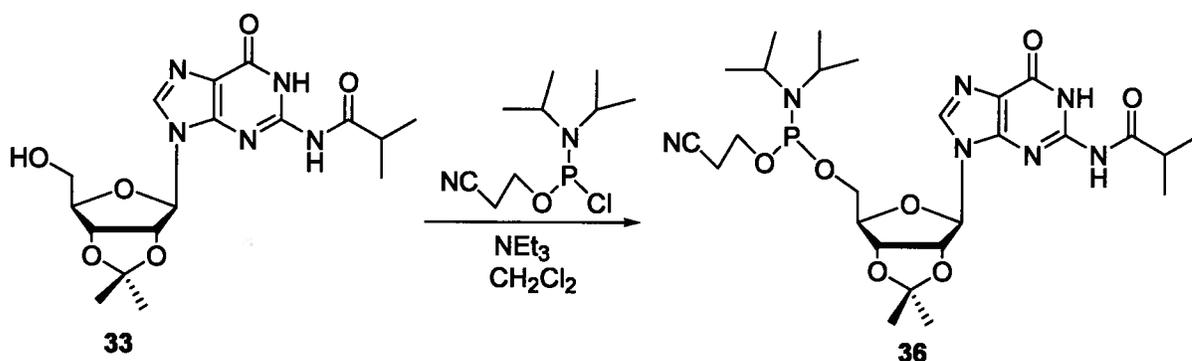
In parallel to the synthesis of 33, methyl-footed aryl tetrol 5a (Scheme 2.2)<sup>51</sup>, triethylamine, and 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite were reacted in dry dichloromethane to form the tetrakisphosphoramidite cavitand 34 (Scheme 4.7). The consumption of the phosphorylating reagent and the formation of the product during the course of the reaction were monitored by <sup>31</sup>P NMR spectroscopy. The TLC and ESI-MS analyses confirmed the complete consumption of the starting material and the formation of the product, but the product turned out to be very labile and hard to purify. This difficulty in purification might be attributed to the formation of anionic tetrol which is a good leaving group under the experimental conditions.

**Scheme 4.7.** The synthesis of tetrakisphosphoramidite cavitand 34.



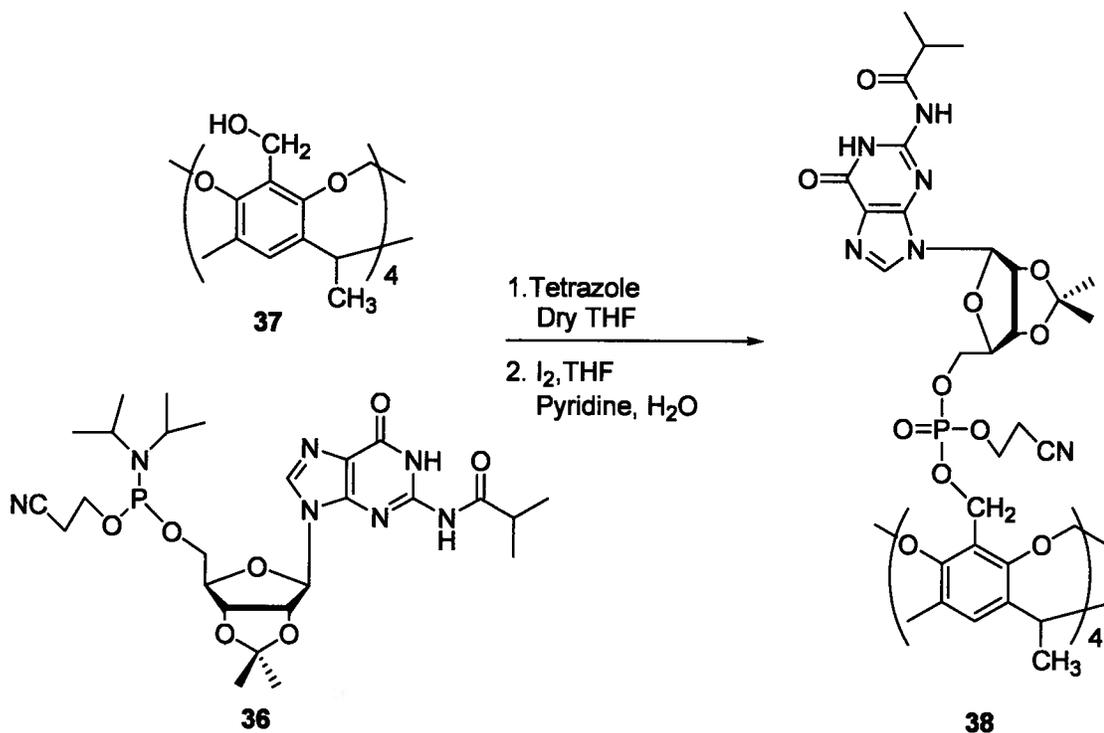
Due to the difficulty in the purification of **34**, we chose to use it without further purification. The amidite **34** was first activated by tetrazole and then the protected guanosine **33** was added in excess. A solution of iodine was then delivered to the solution to oxidize the phosphorus. After a preliminary work-up, the product was analyzed by MALDI-MS but no indication of the presence of the four substituted product was observed. Given the poor results obtained in the preceding reaction, we turned to a new set of reactants. A guanosine phosphoramidite reagent (**36**) was synthesized (Scheme 4.8, 79% yield) which remains to be linked to a benzyl tetrol cavitand (**37**)<sup>52</sup> (Scheme 4.9).

**Scheme 4.8.** The synthesis of *N*<sup>2</sup>-isobutyryl-2',3'-*O*-isopropylidene guanosine phosphoramidite.



This approach is expected to overcome the problems associated with the previous reaction: Benzyl tetrol (pK<sub>a</sub> about 15) is not a good leaving group compared to aryl tetrol (pK<sub>a</sub> about 10) and the guanosine phosphoramidite can be purified by conventional methods. This route was abandoned at this stage to pursue and develop the lipophilic TASQ project but it is worthy of further pursuit (Grant Bare is pursuing it).

**Scheme 4.9.** Proposed synthesis of a hydrophilic TASQ using the phosphoramidite approach.



#### 4.4 Conclusions

Three approaches for the preparation of the hydrophilic TASQs were described in this Chapter. A review of the methods and the progress made in each approach are stated below:

The first approach requires the synthesis of an azide-functionalized oligonucleotide which will be clicked to a methyl-footed cavitand template. Two guanine-rich sequences of 5'-d[G<sub>4</sub>]-3' and 5'-d[TG<sub>4</sub>T]-3' were selected and converted to their corresponding 5'-azido derivatives via a solid-phase synthesis. The preliminary coupling reactions between the cavitand and 5'-d[G<sub>4</sub>]-3' did not lead to the formation of the target product. A relatively wide range of

conditions are still left to be investigated. It is hoped that further exploration of the coupling conditions will lead to the synthesis of the hydrophilic TASQs.

The second approach is similar to the first, except that it takes advantage of a water-soluble cavitand which can be clicked to a lipophilic or hydrophilic guanosine compound. A water-soluble cavitand containing propargyl ether groups on the upper rim and hydrophilic phosphate groups on the lower rim was designed and synthesized for this purpose. The final coupling reaction between the cavitand and guanosine segments showed a promising  $^1\text{H}$  NMR spectrum, but no MALDI-MS signal could be detected for the product. Additional experiments is required to make the final conclusion

The last method utilized for the preparation of the hydrophilic TASQs is based on the phosphoramidite chemistry. An *N*-protected guanosine and a tetrakisphosphoramidite cavitand were prepared. The latter reagent was labile and had limited stability. The coupling reactions performed on the crude reagent were not successful. Therefore, a new set of reactants was synthesized and purified. This route can be followed in parallel or as an alternative to the “click” strategy.

In summary, the reactants needed for the synthesis of the hydrophilic TASQs in all three methods have been successfully synthesized, while the coupling reactions remain to be further researched. The “click” strategy seems to be more promising due to its convenience, atom efficiency and also obviating the need for the protection and deprotection steps. The next Chapter will provide a summary of the current research and suggest several directions which are worthy of consideration for future research.

## 4.5. Experimental

### 4.5.1 Phosphate-Footed Tetrakis (*O*-Propargyl) Cavitand Synthesis

The synthesis of cavitand **30** was accomplished in 9 steps. Compounds **22-24** have been reported in the literature.<sup>42</sup> Compounds **25-30** have not been previously reported.

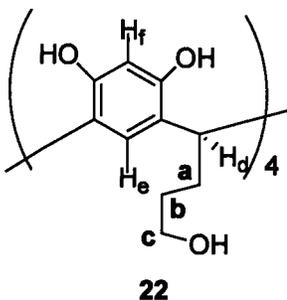
#### 4.5.1.1 General

All chemicals were reagent grade (Sigma-Aldrich) except di-*tert*-butyl *N,N*-diethylphosphoramidite (Toronto Research Chemicals) which was technical grade and used without further purification. Reagent grade THF was distilled under nitrogen from sodium benzophenone ketyl. *N,N*-dimethylformamide (DMF) was dried over 4 Å molecular sieves. Column chromatography was used for purification using silica gel (230-400 mesh, BDH) and silical gel glass backed analytical plates (0.2 mm) were used for thin layer chromatography (TLC) with UV detection.

Deuterated solvents were purchased from Cambridge Isotope Laboratories. <sup>1</sup>H NMR spectra were recorded on Bruker AV-400inv, and AV-400dir spectrometers using the residual solvent signal as the reference. <sup>31</sup>P NMR spectroscopic experiments were performed on a Bruker AV-400inv spectrometer referenced to external H<sub>3</sub>PO<sub>4</sub> (0.00) ppm. An Esquire~LC ion trap mass spectrometer equipped with an electrospray ion source was utilized for the ESI-MS experiments.

### 4.5.1.2 Synthesis of Dodecol 22

Dodecol **22** was prepared through a known procedure<sup>42</sup> and the characterization data matched the literature values. A slight modification to the literature procedure was made in the initial step, which involves addition of 2,3-dihydrofuran to a solution of resorcinol in HCl/CH<sub>3</sub>OH. The reagent was added drop wise over a 2 h period. The mixture was then stirred at RT for 4 h, and then was heated to 50 °C.



**Compound 22:** yield 62 %, <sup>1</sup>NMR (400 MHz, DMSO-d<sub>6</sub>, 298 K) δ 8.88 (s, 8H, phenolic OH), 7.22 (s, 4H, H<sub>f</sub>), 6.13 (s, 4H, H<sub>e</sub>), 4.29 (t, 4H, H<sub>d</sub>), 4.18 (t, 4H, OH), 3.40 (m, 8H, CH<sub>2c</sub>), 2.08 (m, 8H, CH<sub>2b</sub>), 1.33 (m, 8H, CH<sub>2a</sub>) ppm.

**MS (ESI) *m/z*:** calculated: 743.31 for C<sub>40</sub>H<sub>48</sub>O<sub>12</sub>·Na; found: [M+Na<sup>+</sup>] 743.2.

### 4.5.1.3 Synthesis of Tetrabromo Dodecol 23

Tetrabromo dodecol **23** was synthesized following the literature procedure.<sup>42</sup>

**Compound 23:** yield 80 %,  $^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ , 298 K)  $\delta$  9.11 (s, 8H, phenolic OH), 7.40 (s, 4H,  $H_e$ ), 4.34 (t, 4H,  $H_d$ ), 3.60 (br, 4H, OH), 3.43 (m, 8H,  $\text{CH}_{2c}$ ), 2.21 (m, 8H,  $\text{CH}_{2b}$ ), 1.33 (m, 8H,  $\text{CH}_{2a}$ ) ppm.

**MS (ESI)  $m/z$ :** calculated: 1058.95 for  $\text{C}_{40}\text{H}_{44}\text{O}_{12}\text{Br}_4\cdot\text{Na}$ ; found:  $[\text{M}+\text{Na}^+]$  1058.8.

#### 4.5.1.4 Synthesis of Hydroxyl-Footed Tetrabromo Cavitand 24

Selective bridging of compound **23** was carried out following the method developed in our group,<sup>42</sup> but with some modifications in the stoichiometries and work up stage. Tetrabromo dodecol **23** (10.36 g, 10 mmol) and potassium carbonate (40 g, 288 mmol) were added to 400 mL of DMA. After stirring at room temperature for 10 minutes, 30 mL (450 mmol) of bromochloromethane was added and the solution mixture was heated to 65 °C. An additional 5 mL (77.5 mmol) of bromochloromethane was added after each 24 h period and the reaction mixture was stirred at this temperature for 4 days. The solution was then cooled, and the solvent was evaporated *in vacuo*. The solid residue was sonicated in water and the carbonate salt was carefully neutralized by the addition of 2 M HCl. The resulting mixture was filtered and the crude precipitate was washed with water, dissolved in THF and dried over  $\text{MgSO}_4$ . The solvent was evaporated *in vacuo* and the crude reaction mixture was embedded on silica gel. The by-products which had slightly higher  $R_f$  values were first eluted with a mixture of 9:1  $\text{CHCl}_3$ :EtOH (v/v). Then, the pure product was collected using 9:1  $\text{CHCl}_3$ :MeOH (v/v).

**Compound 24:** yield 47 %,  $^1\text{NMR}$  (400 MHz, DMSO- $d_6$ , 298 K)  $\delta$  7.65 (s, 4H, H<sub>e</sub>), 5.99 (d, 4H, H<sub>out</sub>), 4.70 (t, 4H, H<sub>d</sub>), 4.50 (t, 4H, OH), 4.30 (d, 4H, H<sub>in</sub>), 3.50 (m, 8H, CH<sub>2c</sub>), 2.46 (m, 8H, CH<sub>2b</sub>), 1.44 (m, 8H, CH<sub>2a</sub>) ppm.

**MS (ESI)  $m/z$ :** calculated: 1106.95 for C<sub>44</sub>H<sub>44</sub>O<sub>12</sub>Br<sub>4</sub>·Na; found: [M+Na<sup>+</sup>] 1106.8.

#### 4.5.1.5 Synthesis of TBDMS-Protected Tetrabromo Cavitand 25

TBDMS-Cl (4.1 g, 27.3 mmol) was added to a solution of cavitand **24** (3 g, 2.77 mmol) and imidazole (3.7 g, 54.3) in 75 mL of dry DMF and stirred under N<sub>2</sub> at RT. Additional TBDMS-Cl (2 g, 13.3 mmol) was added to this solution after 24 h and the reaction mixture was stirred for one more day at RT. The solution was then slowly poured into chloroform (150 mL) and extracted with water. The aqueous phase was washed twice with chloroform (2×30 mL) and the organic layers were combined and dried over MgSO<sub>4</sub>. The solvent was evaporated *in vacuo* and the crude product was purified by column chromatography on silica gel using 95:5 CH<sub>2</sub>Cl<sub>2</sub>:MeOH (v/v).

**Compound 25:** yield 67%,  $^1\text{NMR}$  (400 MHz, CDCl<sub>3</sub>, 298 K)  $\delta$  7.04 (s, 4H, H<sub>e</sub>), 5.95 (d, 4H, H<sub>out</sub>), 4.88 (t, 4H, H<sub>d</sub>), 4.39 (d, 4H, H<sub>in</sub>), 3.70 (t, 8H, CH<sub>2c</sub>), 2.26 (m, 8H, CH<sub>2b</sub>), 1.56 (m, 8H, CH<sub>2a</sub>), 0.93 (s, 36H, Si-tBu), 0.08 (d, 24H, Si-CH<sub>3</sub>) ppm.

**MS (ESI)  $m/z$ :** calculated: 1563.30 for C<sub>68</sub>H<sub>100</sub>O<sub>12</sub>Br<sub>4</sub>Si<sub>4</sub>·Na; found: [M+Na<sup>+</sup>] 1563.5.

#### 4.5.1.6 Synthesis of TBDMS-Protected Tetrol 26

Cavitand **25** (1.35 g, 0.87 mmol) which had been dried *in vacuo* for 2 days was placed in a round bottom flask and further dried by coevaporation with THF (250 mL). Then, 210 mL of dry THF was delivered into the flask and the solution was cooled to  $-78\text{ }^{\circ}\text{C}$ . *n*-BuLi (5 mL of a 1.6 M solution in hexanes, 7.5 mmol) was added to this solution followed by the addition of  $\text{B}(\text{OMe})_3$  (1 mL, 4.3 mmol) and the solution was allowed to warm to room temperature over 2 h. The reaction mixture was again cooled to  $-78\text{ }^{\circ}\text{C}$ , 1.5 M NaOH-15%  $\text{H}_2\text{O}_2$  (23 mL) was added and the reaction mixture was allowed to warm to room temperature over 2 h. The excess  $\text{H}_2\text{O}_2$  was neutralized by the careful addition of  $\text{Na}_2\text{S}_2\text{O}_5$  (3.8 g, 20 mmol). The solvent was then evaporated *in vacuo* and the solid residue was filtered and washed with water. The crude product was purified by column chromatography on silica gel using 10:1 EtOAc:EtOH.

**Compound 26:** yield 31%,  $^1\text{NMR}$  (400 MHz,  $\text{CDCl}_3$ , 298 K)  $\delta$  6.64 (s, 4H,  $\text{H}_e$ ), 5.96 (d, 4H,  $\text{H}_{\text{out}}$ ), 5.35 (br, 4H, phenolic OH), 4.73 (t, 4H,  $\text{H}_d$ ), 4.44 (d, 4H,  $\text{H}_{\text{in}}$ ), 3.70 (t, 8H,  $\text{CH}_{2c}$ ), 2.22 (m, 8H,  $\text{CH}_{2b}$ ), 1.58 (m, 8H,  $\text{CH}_{2a}$ ), 0.93 (s, 36H, Si-*t*Bu), 0.08 (d, 24H, Si- $\text{CH}_3$ ) ppm.

**MS (ESI) *m/z*:** calculated: 1311.64 for  $\text{C}_{68}\text{H}_{104}\text{O}_{16}\text{Si}_4\cdot\text{Na}$ ; found:  $[\text{M}+\text{Na}^+]$  1311.5.

#### 4.5.1.7 Synthesis of TBDMS-Protected Tetrakis (*O*-Propargyl) Cavitand 27

Tetrol **26** (0.47 mg, 0.36 mmol), potassium carbonate (1 g, 7.3 mmol) and 0.60 mL 80% propargyl bromide in toluene (5.3 mmol) were added to 25 mL of acetone and were heated to

reflux for 20 h. The solvent was evaporated *in vacuo* and the residue was taken up in 200 mL of CH<sub>2</sub>Cl<sub>2</sub> and stirred. The resulting suspension was filtered and the filtrate was dried under reduced pressure and subjected to column chromatography on silica gel. The pure product was eluted with 95:5 CH<sub>2</sub>Cl<sub>2</sub>:EtOH.

**Compound 27:** yield 72%, <sup>1</sup>NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 6.81 (s, 4H, H<sub>e</sub>), 5.89 (d, 4H, H<sub>out</sub>), 4.73 (t, 4H, H<sub>d</sub>), 4.59 (d, 8H, O-CH<sub>2</sub> propargyl), 4.39 (d, 4H, H<sub>in</sub>), 3.70 (t, 8H, CH<sub>2c</sub>), 2.47 (t, 4H, C≡CH), 2.22 (m, 8H, CH<sub>2b</sub>), 1.58 (m, 8H, CH<sub>2a</sub>), 0.93 (s, 36H, Si-tBu), 0.08 (d, 24H, Si-CH<sub>3</sub>) ppm.

**MS (ESI) *m/z*:** calculated: 1463.70 for C<sub>80</sub>H<sub>112</sub>O<sub>16</sub>Si<sub>4</sub>·Na; found: [M+Na<sup>+</sup>] 1463.5.

#### 4.5.1.8 Synthesis of Hydroxyl-Footed Tetrakis (*O*-Propargyl) Cavitand 28

Anhydrous tetrabutylammonium fluoride (1.2 mL of a 1.0 M solution in THF, 1.2 mmol) was added to a solution of cavitand 27 (200 mg, 0.138 mmol) in dry THF (8 mL) at 0 °C. The solution was then warmed to room temperature and stirred at this temperature for 40 minutes. The solvent was evaporated *in vacuo* and the residue was embedded on silica gel and subjected to column chromatography on silica gel. The pure product was collected using 5:0.5 EtOAc:EtOH.

**Compound 28:** yield 60%, <sup>1</sup>NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 6.95 (s, 4H, H<sub>e</sub>), 5.90 (d, 4H, H<sub>out</sub>), 4.75 (t, 4H, H<sub>d</sub>), 4.61 (d, 8H, O-CH<sub>2</sub><sub>propargyl</sub>), 4.41 (d, 4H, H<sub>in</sub>), 3.79 (t, 8H, CH<sub>2c</sub>), 3.34 (br, 4H, OH), 2.47 (t, 4H, C≡CH), 2.36 (m, 8H, CH<sub>2b</sub>), 1.62 (m, 8H, CH<sub>2a</sub>), ppm.

**MS (ESI) *m/z*:** calculated: 1007.36 for C<sub>56</sub>H<sub>56</sub>O<sub>16</sub>·Na; found: [M+Na<sup>+</sup>] 1007.3.

#### 4.5.1.9 Synthesis of t-Butyl Phosphorylated Tetrakis (*O*-Propargyl) Cavitand 29

Di-tert-butyl *N,N*-diethylphosphoramidite (0.300 mL, 1.06 mmol) was added to a suspension of 1-H-tetrazole (225 mg, 3.19 mmol) and cavitand **28** (70 mg, 0.071 mmol) in dry THF and the solution was stirred under N<sub>2</sub> for 20 minutes. The reaction mixture was cooled to -78 °C and H<sub>2</sub>O<sub>2</sub> (0.150 mL of a 30% solution, 1.37 mmol) was added. After warming to room temperature over 30 minutes, the solution was poured into water and extracted with CHCl<sub>3</sub> (2× 50 mL). The organic phase was separated and washed with 10% NaHSO<sub>3</sub> and water and dried over MgSO<sub>4</sub>. The solvent was evaporated and the residue was subjected to column chromatography on silica gel using 95:5 CHCl<sub>3</sub>:MeOH.

**Compound 29:** yield 61%, <sup>1</sup>NMR (400 MHz, CDCl<sub>3</sub>, 298 K) δ 6.81 (s, 4H, H<sub>e</sub>), 5.89 (d, 4H, H<sub>out</sub>), 4.77 (t, 4H, H<sub>d</sub>), 4.59 (d, 8H, O-CH<sub>2</sub><sub>propargyl</sub>), 4.38 (d, 4H, H<sub>in</sub>), 4.03 (m, 8H, CH<sub>2c</sub>), 2.48 (t, 4H, C≡CH), 2.32 (m, 8H, CH<sub>2b</sub>), 1.71 (m, 8H, CH<sub>2a</sub>), 1.48 (s, 72H, tBu) ppm.

**MS (ESI) *m/z*:** calculated: 1776.73 for C<sub>88</sub>H<sub>124</sub>O<sub>28</sub>P<sub>4</sub>·Na; found: [M+Na<sup>+</sup>] 1776.6.

#### 4.5.1.10 Synthesis of Phosphate-Footed Tetrakis (*O*-Propargyl) Cavitand **30**

Trifluoroacetic acid (0.32 mL, 4.2 mmol) was added to a solution of cavitand **29** (20 mg, 0.011 mmol) in CH<sub>2</sub>Cl<sub>2</sub> and stirred for 30 minutes. The pure product was collected after evaporation of the solvent.

**Compound 30:** yield 97%, <sup>1</sup>NMR (400 MHz, CD<sub>3</sub>OD, 298 K) δ 7.1 (s, 4H, H<sub>e</sub>), 5.86 (d, 4H, H<sub>out</sub>), 4.70 (t, 4H, H<sub>d</sub>), 4.61 (d, 8H, O-CH<sub>2</sub> propargyl), 4.38 (d, 4H, H<sub>in</sub>), 4.11 (m, 8H, CH<sub>2c</sub>), 2.89 (t, 4H, C≡CH), 2.42 (m, 8H, CH<sub>2b</sub>), 1.69 (m, 8H, CH<sub>2a</sub>) ppm.

<sup>31</sup>P NMR: (CD<sub>3</sub>OD, 81 MHz, 298 K) δ 1.0 (s, 4P) ppm.

MS (ESI) *m/z*: calculated: 1304.22 for C<sub>56</sub>H<sub>60</sub>O<sub>28</sub>P<sub>4</sub>; found: [M-H]<sup>-</sup> 1303.1.

## 4.5.2 Oligonucleotide Synthesis

### 4.5.2.1 General

All chemicals were purchased from Sigma-Aldrich Chemical Co. Inc., except the oligonucleotides which were obtained from the DNA/RNA synthesis laboratory of the University of Calgary (Faculty of Medicine). The oligonucleotides had all been synthesized using standard phosphoramidite method and were received on CPG support with the DMT group off. The empty DNA synthesizing columns were purchased from 3-prime Inc. (ABI snap style 40  $\mu$ mole). HPLC experiments were conducted on a Waters Delta 600 system equipped with a Waters 2996 photodiode array (UV) detector. The wavelength for the UV detection was set at 260 nm. Both analytical and preparative HPLC were run on Phenomenex Clarity C<sub>18</sub> reversed-phase columns (preparative: 100 mm  $\times$  10 mm, 5  $\mu$ M particle size, and analytical: 100 mm  $\times$  4.6 mm, 3  $\mu$ M particle size). The preparative samples were filtered through a 0.45  $\mu$ M Nylon™ syringe filter (Phenomenex) prior to injection and run at a flow rate of 4 mL/min using triethylammonium acetate buffer (50 mM)/HPLC-grade acetonitrile (0.05 % TFA) gradient. The analytical samples were run under similar conditions with the flow rate of 1 mL/min. Concentrations of the samples were determined using UV-Vis measurements obtained on a Varian Cary 3E spectrophotometer. The purified samples were concentrated *in vacuo* and lyophilized. The mass spectra were run on a MALDI-TOF Bruker Biflex IV in reflectron mode using a 3-hydroxypicolinic acid (3-HPA):imidazole matrix:co-matrix combination.

#### 4.5.2.2 Synthesis of 5'-Iodo Oligonucleotides 18a-b

The iodination of the oligonucleotides were performed following the general procedure described by Kool and co-workers.<sup>25</sup> Two oligonucleotides of the sequences of 5'-d[G<sub>4</sub>]-3' (**17a**) and 5'-d[TG<sub>4</sub>T]-3' (**17b**) were selected and subjected to the modification. In a typical experiment, 0.24 g of the CPG resin (containing approximately 3.24  $\mu$ mol of the oligonucleotide) was loaded into an ABI snap style DNA synthesizing column (40  $\mu$ mole) and rinsed with 2 mL of dry DMF. Two 1-mL syringes were then attached to this column and a solution of methyltriphenoxyphosphonium iodide (1.0 mL of a 0.5 M solution in DMF) was passed through it for 1 minute. The syringes were removed and the column was sealed with two plugs and placed on a shaker at room temperature for 15 minutes. The reagent was emptied from the column and the CPG resin was washed with 10 mL of CH<sub>2</sub>Cl<sub>2</sub>, 5 mL of CH<sub>3</sub>CN, and 5 mL of CH<sub>2</sub>Cl<sub>2</sub>. The 5'-iodinated oligonucleotides **18a-b** were not cleaved from the solid support, but were directly converted to the corresponding azides.<sup>24</sup>

#### 4.5.2.3 Synthesis of 5'-Azido Oligonucleotides 20a-b

A DNA synthesizing column containing the iodinated product of the previous step (**18a** or **18b**) was washed with 2 mL of dry DMF to remove the absorbed water. The column was then fitted with two 1-mL syringes and a saturated solution of sodium azide in DMF was passed through it for 1 minute. The syringes were removed and the column was sealed and heated at 60 °C for 1 h. The reagent was emptied from the column and the CPG resin was washed with 10 mL of DMF and 5 mL of CH<sub>2</sub>Cl<sub>2</sub> and then blown dry by air. The oligonucleotide was then

cleaved from the solid support using a fresh ammonium hydroxide solution (28-30%) at room temperature for 2 h. The oligonucleotide which was suspended in this solution was sealed in closed vials and heated for 8 h at 55 °C to deprotect the bases. The vials were then removed from the heating blocks and put into the fridge for 30 minutes. Once cool, they were placed into a speed-vac to dry with moderate heating (under the vacuum). The resulting pellets were each dissolved in 100  $\mu$ L of concentrated ammonium hydroxide solution (30%) and vigorously vortexed. To this solution was added 1 mL of n-butanol and the mixture was vortexed again. The vigorous vortex continued until the oligonucleotides precipitated out of the solution. This cloudy solution was centrifuged and the resulting pellets were carefully separated from the supernatants. The pellets were dissolved in water, combined and lyophilized.

An analytical sample was purified by HPLC and the identities of the products were confirmed using mass-spectroscopic techniques (MALDI-TOF-MS and ESI-MS) and observation of a single peak by analytical reversed-phase HPLC (Figures 4.3, 4.4, 4.5, and 4.6). The yields reported here are relative yields obtained from inspection of the HPLC chromatograms. The reaction works equally well with the high loading resin,  $\sim$ 56  $\mu$ mol/g.

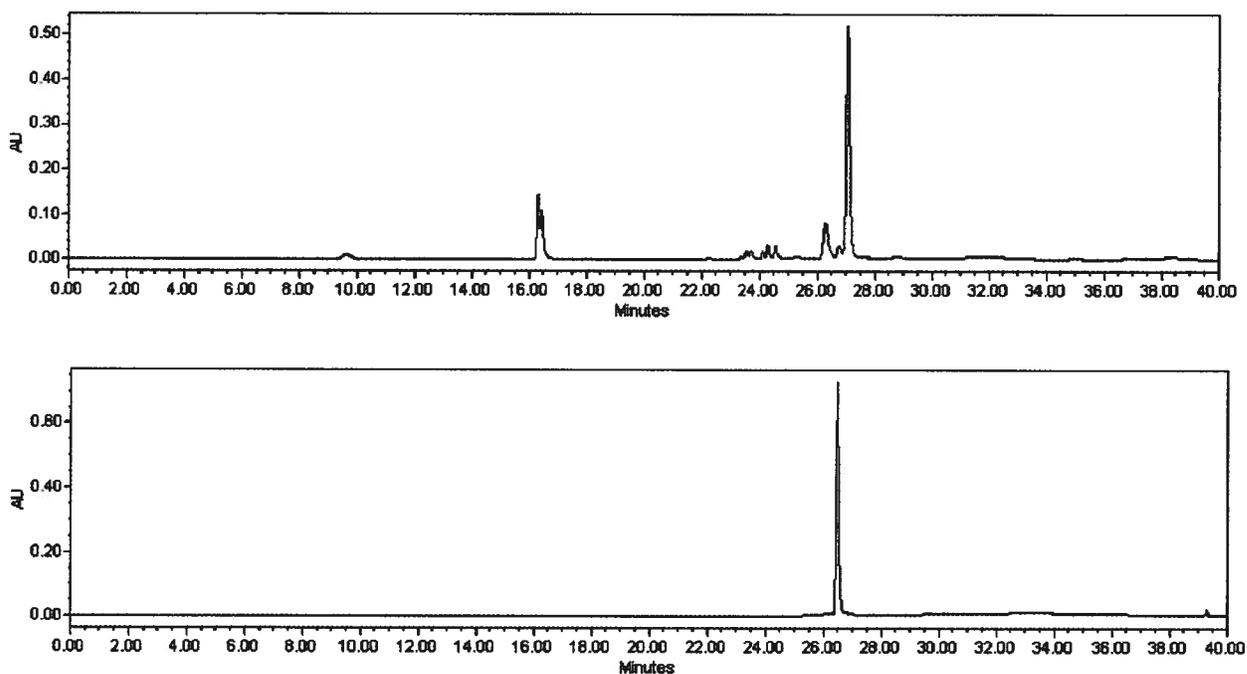
**Compound 20a (5'- d[N<sub>3</sub>-G<sub>4</sub>]-3')**: yield (5'-OH to 5'-N<sub>3</sub>), 60-70%

**MS (ESI) *m/z***: calculated: 1279.26 for C<sub>40</sub>H<sub>48</sub>N<sub>23</sub>O<sub>21</sub>P<sub>3</sub>; found: [M-H]<sup>-</sup> 1278.2

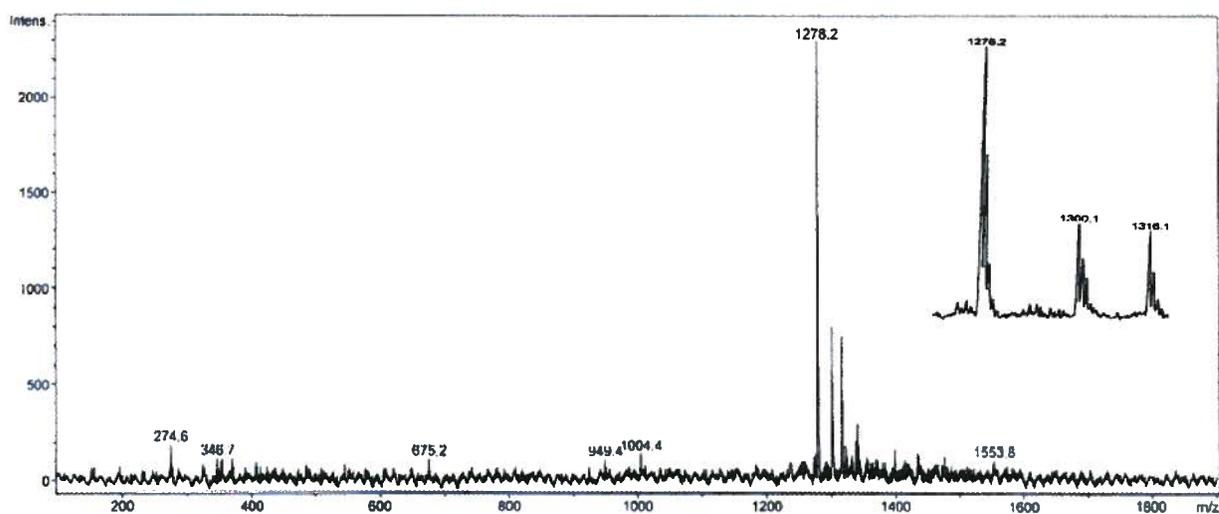
**Compound 20b (5'- d[N<sub>3</sub>-TG<sub>4</sub>T]-3')**: yield (5'-OH to 5'-N<sub>3</sub>), 70-80%

**MS (MALDI-TOF) *m/z***: calculated: 1887.35 for C<sub>60</sub>H<sub>74</sub>N<sub>27</sub>O<sub>35</sub>P<sub>5</sub>; found: [M+H]<sup>+</sup> 1888.8

**Figure 4.3.** HPLC chromatograms of the crude (top), and the pure **20a** (bottom).



**Figure 4.4.** ESI-MS spectrum of **20a** in negative mode. The signals at ( $m/z = 1278.2$ ,  $1300.1$ , and  $1316.1$ ) correspond to  $[M-H]^-$ ,  $[M+Na-2H]^-$ , and  $[M+K-2H]^-$  respectively.





### 4.5.3 Phosphoramidite Synthesis

#### 4.5.3.1 General

See section 4.5.1.1 for the reagents and instrumentation. 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite was obtained from Alfa Aesar.  $^{31}\text{P}$  NMR spectroscopic experiments were performed on a Bruker AC-200 spectrometer referenced to external  $\text{H}_3\text{PO}_4$  (0.00) ppm.

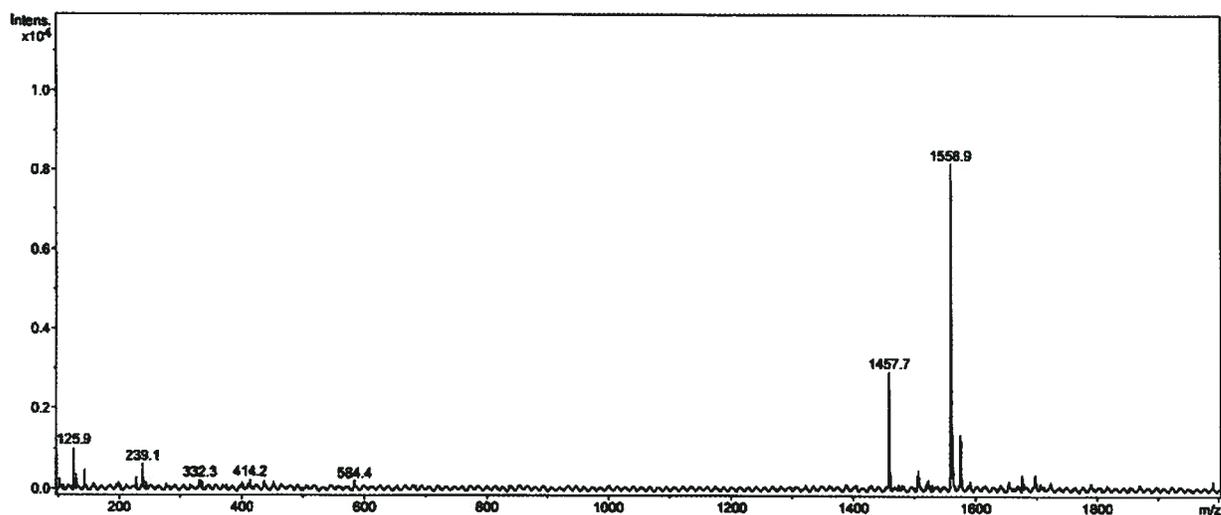
#### 4.5.3.2 Synthesis of Tetrakisphosphoramidite Cavitand 34

Tetrol **5a**<sup>51</sup> (20 mg, 0.03 mmol) dried *in vacuo* was suspended in 3 mL of dry  $\text{CH}_2\text{Cl}_2$ . To this solution was added triethylamine (40  $\mu\text{L}$ , 0.28 mmol) and 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite (48  $\mu\text{L}$ , 0.21 mmol). The reaction mixture was stirred under  $\text{N}_2$  for 2 h. The reaction progress was monitored by  $^{31}\text{P}$  NMR spectroscopy (Figure 4.8) and the formation of the product was confirmed by ESI-MS (Figure 4.7) and observation of one TLC spot. This product was labile and could not be isolated.

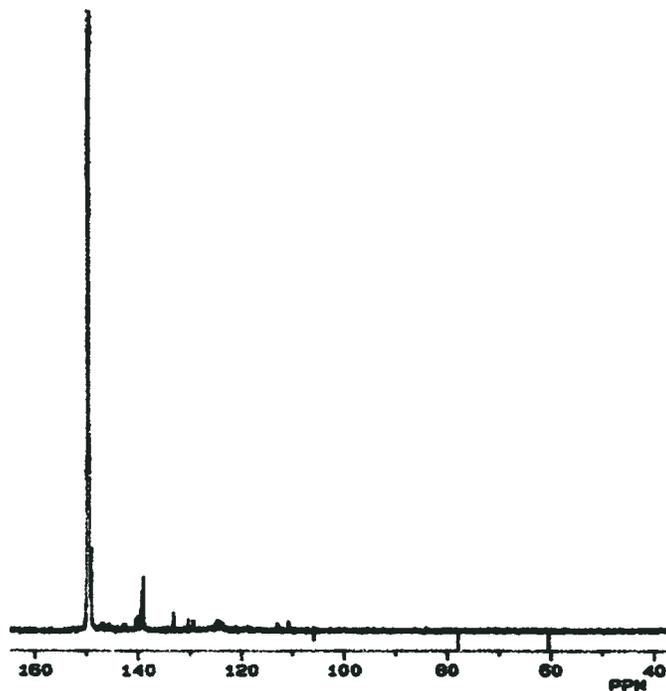
$^{31}\text{P}$  NMR: ( $\text{CD}_2\text{Cl}_2$ , 81 MHz, 298 K)  $\delta$  149.9 (s, 4P) ppm.

MS (ESI) *m/z*: calculated: 1456.62 for  $\text{C}_{72}\text{H}_{100}\text{N}_8\text{O}_{16}\text{P}_4$ ; found:  $[\text{M}+\text{H}^+]$  1457.7, and  $[\text{M}+\text{NEt}_3+\text{H}^+]$  1558.9.

**Figure 4.7.** ESI-MS spectrum of **34** in positive mode. Only one species was observed in the solution. The signals at ( $m/z = 1457.7$ ) and ( $m/z = 1558.9$ ) correspond to  $[M+H]^+$  and  $[M+NEt_3+H]^+$  respectively.



**Figure 4.8.** The 81 MHz  $^{31}P$  NMR spectrum of a solution of **5a** reacted with 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite in  $CD_2Cl_2$  at 25 °C. The signal at  $\sim 149.9$  ppm corresponds to the product. (Note that the P-diastereomers are not resolved at 81 MHz)



#### 4.5.3.3 Synthesis of *N*<sup>2</sup>-Isobutyryl-2',3'-*O*-Isopropylidene Guanosine 33

The *N*-protected guanosine **33** was prepared following the Jones's transient protection method.<sup>50,53</sup> 2',3'-*O*-isopropylidene guanosine **7** (3.23g, 10 mmol) dried by coevaporation with pyridine (2×50 mL) was suspended in dry pyridine (60 mL) and treated with 3.16 mL (25 mmol) of trimethylchlorosilane. After stirring at room temperature under N<sub>2</sub> for 2 h, 2.90 mL (30 mmol) of isobutyryl chloride was added drop wise and stirring was continued for 3 h. The reaction mixture was cooled in an ice-bath and 10 mL of water was added. After 10 minutes of stirring at room temperature, concentrated ammonia solution (20 mL, 30%) was added and the solution was stirred for 20 minutes. The solvent was evaporated *in vacuo* (under moderate heat) and the residue was partitioned between water and chloroform. The organic layer was separated, dried over MgSO<sub>4</sub> and subjected to column chromatography on silica gel using 95:5 CH<sub>2</sub>Cl<sub>2</sub>: MeOH.

**Compound 33:** yield 71%, <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 12.14 (s, 1H, NH<sub>imino</sub>), 9.20 (s, 1H, NH<sub>amide</sub>), 7.88 (s, 1H, H8), 5.85 (d, 1H, H1'), 5.60 (1H, OH), 5.12 (dd, 1H, H2'), 5.00 (dd, 1H, H3'), 4.40 (m, 1H, H4'), 3.93 (dd, 1H, H5'), 3.78 (dd, 1H, H5'), 2.70 (m, 1H, CH), 1.58 (s, 3H, CH<sub>3</sub>), 1.34 (s, 3H, CH<sub>3</sub>), 1.27 (d, 3H, CH<sub>3</sub>), 1.25 (d, 3H, CH<sub>3</sub>) ppm.

**MS (ESI) *m/z*:** calculated: 393.16 for C<sub>17</sub>H<sub>23</sub>N<sub>5</sub>O<sub>6</sub>; found: [M+H<sup>+</sup>] 394.1.

#### 4.5.3.4 Synthesis of *N*<sup>2</sup>-Isobutyryl-2',3'-*O*-Isopropylidene Guanosine Phosphoramidite 36

To a solution of *N*-protected guanosine **33** (100 mg, 0.25 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (5 mL) was added triethylamine (71 μL, 0.50 mmol) and 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite (85 μL, 0.37 mmol). The solution was stirred under N<sub>2</sub> for 2 h. Then the reaction mixture was partitioned between water and chloroform. The organic layer was separated, dried over MgSO<sub>4</sub> and subjected to column chromatography on silica gel using 95:5 MeOH:CH<sub>2</sub>Cl<sub>2</sub>. The product is a mixture of two diastereomers due to the stereogenic nature of the phosphorous atom (yield 79%).

<sup>31</sup>P NMR: (CH<sub>2</sub>Cl<sub>2</sub>, 81 MHz, 298 K) δ 151.0, 151.3 (d, 1P) ppm.

MS (ESI) *m/z*: calculated: 593.27 for C<sub>26</sub>H<sub>40</sub>N<sub>7</sub>O<sub>7</sub>P; found: [M+H<sup>+</sup>] 594.3.

## 4.6 References

- (1) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-471.
- (2) Li, J. L.; Harrison, R. J.; Reszka, A. P.; Brosh, R. M.; Bohr, V. A.; Neidle, S.; Hickson, I. D. *Biochemistry* **2001**, *40*, 15194-15202.
- (3) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593-11598.
- (4) Sundquist, W. I.; Heaphy, S. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3393-3397.
- (5) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.
- (6) Stewart, S. A.; Weinberg, R. A. *Semin. Cancer Biol.* **2000**, *10*, 399-406.
- (7) Bodnar, A. G.; Ouellette, M.; Frolkis, M.; Holt, S. E.; Chiu, C. P.; Morin, G. B.; Harley, C. B.; Shay, J. W.; Lichtsteiner, S.; Wright, W. E. *Science* **1998**, *279*, 349-352.
- (8) Ghosal, G.; Muniyappa, K. *Biochem. Biophys. Res. Commun.* **2006**, *343*, 1-7.
- (9) Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. *Tetrahedron Lett.* **2004**, *45*, 4869-4872.
- (10) Koizumi, M.; Akahori, K.; Ohmine, T.; Tsutsumi, S.; Sone, J.; Kosaka, T.; Kaneko, M.; Kimura, S.; Shimada, K. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2213-2216.
- (11) Wyatt, J. R.; Vickers, T. A.; Roberson, J. L.; Buckheit, R. W.; Klimkait, T.; Debaets, E.; Davis, P. W.; Rayner, B.; Imbach, J. L.; Ecker, D. J. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 1356-1360.
- (12) Okazawa, A.; Maeda, H.; Fukusaki, E.; Katakura, Y.; Kobayashi, A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2653-2656.
- (13) Gavathiotis, E.; Heald, R. A.; Stevens, M. F. G.; Searle, M. S. *J. Mol. Biol.* **2003**, *334*, 25-36.
- (14) Wyatt, J. R.; Davis, P. W.; Freier, S. M. *Biochemistry* **1996**, *35*, 8002-8008.
- (15) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596-2599.
- (16) Barton, D.; Nakanishi, K. o.; Meth-Cohn, O. *Comprehensive natural products chemistry*; 1st ed.; Elsevier: Amsterdam; New York, 1999; Vol. 7.
- (17) Sen, D.; Gilbert, W. *Biochemistry* **1992**, *31*, 65-70.

- (18) Krishnan-Ghosh, Y.; Liu, D. S.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 11009-11016.
- (19) Kato, Y.; Ohyama, T.; Mita, H.; Yamamoto, Y. *J. Am. Chem. Soc.* **2005**, *127*, 9980-9981.
- (20) Mezo, A. R.; Sherman, J. C. *J. Am. Chem. Soc.* **1999**, *121*, 8983-8994.
- (21) Seo, E. S.; Scott, W. R. P.; Straus, S. K.; Sherman, J. C. *Chem.-Eur. J.* **2007**, *13*, 3596-3605.
- (22) Kohn, M.; Breinbauer, R. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 3106-3116.
- (23) Alvira, M.; Eritja, R. *Chem. Biodivers.* **2007**, *4*, 2798-2809.
- (24) Miller, G. P.; Kool, E. T. *J. Org. Chem.* **2004**, *69*, 2404-2410.
- (25) Miller, G. P.; Kool, E. T. *Org. Lett.* **2002**, *4*, 3599-3601.
- (26) Grimm, G. N.; Boutorine, A. S.; Helene, C. *Nucleosides, Nucleotides Nucleic Acids* **2000**, *19*, 1943-1965.
- (27) Chiou, S. H. *J. Biochem. (Tokyo, Jpn.)* **1983**, *94*, 1259-1267.
- (28) Weller, R. L.; Rajski, S. R. *Org. Lett.* **2005**, *7*, 2141-2144.
- (29) Devaraj, N. K.; Miller, G. P.; Ebina, W.; Kakaradov, B.; Collman, J. P.; Kool, E. T.; Chidsey, C. E. D. *J. Am. Chem. Soc.* **2005**, *127*, 8600-8601.
- (30) Kanan, M. W.; Rozenman, M. M.; Sakurai, K.; Snyder, T. M.; Liu, D. R. *Nature* **2004**, *431*, 545-549.
- (31) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.
- (32) Lipshutz, B. H.; Taft, B. R. *Angew. Chem., Int. Ed.* **2006**, *45*, 8235-8238.
- (33) Chassaing, S.; Sido, A. S. S.; Alix, A.; Kumarraja, M.; Pale, P.; Sommer, J. *Chem.-Eur. J.* **2008**, *14*, 6713-6721.
- (34) Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.* **2004**, *6*, 2853-2855.
- (35) Ryu, E. H.; Zhao, Y. *Org. Lett.* **2005**, *7*, 1035-1037.
- (36) Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *Eur. J. Org. Chem.* **2005**, 51-68.
- (37) Zhang, L.; Chen, X. G.; Xue, P.; Sun, H. H. Y.; Williams, I. D.; Sharpless, K. B.; Fokin, V. V.; Jia, G. C. *J. Am. Chem. Soc.* **2005**, *127*, 15998-15999.

- (38) Golas, P. L.; Tsarevsky, N. V.; Sumerlin, B. S.; Matyjaszewski, K. *Macromolecules* **2006**, *39*, 6451-6457.
- (39) Mezo, A. R.; Sherman, J. C. *J. Org. Chem.* **1998**, *63*, 6824-6829.
- (40) Sebo, L.; Diederich, F.; Gramlich, V. *Helv. Chim. Acta* **2000**, *83*, 93-113.
- (41) Haas, C. H.; Biros, S. M.; Rebek, J. *Chem. Commun.* **2005**, 6044-6045.
- (42) Gibb, B. C.; Chapman, R. G.; Sherman, J. C. *J. Org. Chem.* **1996**, *61*, 1505-1509.
- (43) Consoli, G. M. L.; Granata, G.; Galante, E.; Cunsolo, F.; Geraci, C. *Tetrahedron Lett.* **2006**, *47*, 3245-3249.
- (44) Consoli, G. M. L.; Granata, G.; Galante, E.; Di Silvestro, I.; Salafia, L.; Geraci, C. *Tetrahedron* **2007**, *63*, 10758-10763.
- (45) Consoli, G. M. L.; Granata, G.; Garozzo, D.; Mecca, T.; Geraci, C. *Tetrahedron Lett.* **2007**, *48*, 7974-7977.
- (46) Stewart, K. M.; Rojo, J.; McLaughlin, L. W. *Angew. Chem., Int. Ed.* **2004**, *43*, 5808-5811.
- (47) Stewart, K. M.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2004**, *126*, 2050-2057.
- (48) Vargas-Baca, I.; Mitra, D.; Zulyniak, H. J.; Banerjee, J.; Sleiman, H. F. *Angew. Chem., Int. Ed.* **2001**, *40*, 4629-4632.
- (49) Oliviero, G.; Amato, J.; Borbone, N.; Galeone, A.; Petraccone, L.; Varra, M.; Piccialli, G.; Mayol, L. *Bioconjugate Chem.* **2006**, *17*, 889-898.
- (50) Ti, G. S.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 1316-1319.
- (51) Sherman, J. C.; Knobler, C. B.; Cram, D. J. *J. Am. Chem. Soc.* **1991**, *113*, 2194-2204.
- (52) Moran, J. R.; Karbach, S.; Cram, D. J. *J. Am. Chem. Soc.* **1982**, *104*, 5826-5828.
- (53) Eisenfuhr, A.; Arora, P. S.; Sengle, G.; Takaoka, L. R.; Nowick, J. S.; Famulok, M. *Bioorg. Med. Chem.* **2003**, *11*, 235-249.

## CHAPTER 5: Summary, and Future Work

### 5.1 Thesis Summary

G-quartet chemistry is rapidly expanding as it becomes multi-disciplinary.<sup>1</sup> Since their discovery, G-quartets and G-quadruplexes have been proposed and examined for a wide range of applications including their use as supramolecular polymers,<sup>2</sup> nanomachines,<sup>3</sup> liquid crystalline phases<sup>4</sup> and electronic devices.<sup>5</sup> The identification of G-quadruplexes formed from telomeric DNA and other guanine-rich sequences has led to speculations that these structures play an important role in some biological processes. Telomeric DNA for example suffers end degradation and allows for a limited number of cell divisions. This sequence may be involved in the formation of G-quadruplexes which might influence the function of the telomerase (DNA-extension) enzyme *in vivo*.<sup>1,6,7</sup>

Nucleic acid G-quadruplexes can exist in a variety of stoichiometries and folding patterns.<sup>8</sup> As one of the major goals of this project, we proposed that a cavitand template can potentially control the morphology of G-quadruplex assemblies. Following this, the first class of lipophilic template-assembled synthetic G-quartets (TASQs) was synthesized using the “click”<sup>9</sup> reaction. This method employed a 5'-azido guanosine, which was efficiently coupled to three different tetrapropargyl ether cavitand templates. Initial <sup>1</sup>H NMR spectroscopic analysis showed that these compounds aggregate in chloroform. Further investigations revealed their structural details and demonstrated that TASQs form an all-*syn* cation-free G-quartet in solution. Using PFG-NMR experiments, we determined their stoichiometries and concluded that TASQs form a basket-like unimolecular assembly in chloroform. The analysis of the results obtained from

variable-temperature  $^1\text{H}$  NMR and IR spectroscopic measurements revealed that TASQs are thermodynamically stable and this stability is likely offered by the cavitand template. In addition, the results of CD spectroscopy demonstrated that TASQs denature in DMSO in a manner consistent with our  $^1\text{H}$  NMR observation.

The binding of metal cations to nucleic acid G-quadruplexes has been shown to have important topological and most likely biological consequences.<sup>10</sup> In an effort to study the effect of cations on topology of TASQs, as well as to examine the affinity and selectivity of TASQs for cations, we prepared and characterized the cation-bound complexes  $\text{TASQ}\cdot\text{Na}^+$ ,  $\text{TASQ}\cdot\text{K}^+$ ,  $\text{TASQ}\cdot\text{Sr}^{2+}$  and  $\text{TASQ}\cdot\text{Cs}^+$ . From examinations of these compounds using NMR spectroscopy, we determined that a G-quartet is the basic unit of these assemblies. Diffusion NMR analysis revealed that  $\text{TASQ}\cdot\text{Na}^+$  and  $\text{TASQ}\cdot\text{Sr}^{2+}$  exist as monomers while  $\text{TASQ}\cdot\text{Cs}^+$  forms a dimer and  $\text{TASQ}\cdot\text{K}^+$  is most probably a mixture of a monomer and dimer. In addition, the formation of the monomeric assemblies was found to be controlled by the anion. In the cation-competition experiments, the cation-free TASQ demonstrated a relatively high affinity and selectivity for  $\text{Sr}^{2+}$ . Moreover, compared to nucleotides such as 5'-GMP, TASQs also showed an enhanced affinity for cesium. The overall cation affinity order was evaluated as follow in the presence of picrate anion:  $\text{Sr}^{2+} > \text{Na}^+ > \text{K}^+ > \text{Cs}^+$ .

Following the success in synthesizing the lipophilic TASQs, we explored the possibility of synthesizing a hydrophilic TASQ. Two azide-functionalized oligonucleotides were synthesized and their “click” reaction with the cavitands was investigated under different conditions. In contrast to the efficient and convenient synthesis of the lipophilic TASQs, the “click” reaction of the cavitand and oligonucleotides did not afford the desired products. This led to further investigations to explain the rather unexpected results. A number of factors which might influence the “click” reaction were examined. While some concerns, such as the

difference in the solubility of the reactants, seem to have been answered, others such as optimal loading of the copper catalyst remain to be thoroughly examined. On the basis of the early evidence and available literature, a hypothesis was proposed that the saturation of the copper catalyst might be the factor preventing the reaction of the cavitand. In addition, since the problem has only been observed for the synthesis of the hydrophilic TASQs it would be logical to assume that the deactivation of the catalyst is probably intensified under certain conditions.

In a closely related route we explored the possibility of synthesizing a hydrophilic TASQ using a water-soluble cavitand. The key reactant needed for this reaction, a phosphate-footed tetrapropargyl ether cavitand, was synthesized in 9 steps in small-scale. The preliminary coupling reactions (two attempts) yielded a promising NMR spectrum, but the product was not detected by MALDI. With the synthetic route to the cavitand now established, it is possible to repeat the synthesis and the coupling reactions in a larger scale. This line of research remains open for investigation and along with the work described in the previous paragraph; the extension of the lipophilic project (section 5.2.1) forms the basis of the future research.

## **5.2 Future Work**

The work in this thesis offered new development towards synthetic modification and topological control of G-quadruplexes. It also provided a foundation upon which more sophisticated structures can be developed. For future research, the following areas of study would be interesting to pursue.

## 5.2.1 Lipophilic TASQs Project

### 5.2.1.1 Complementary Work

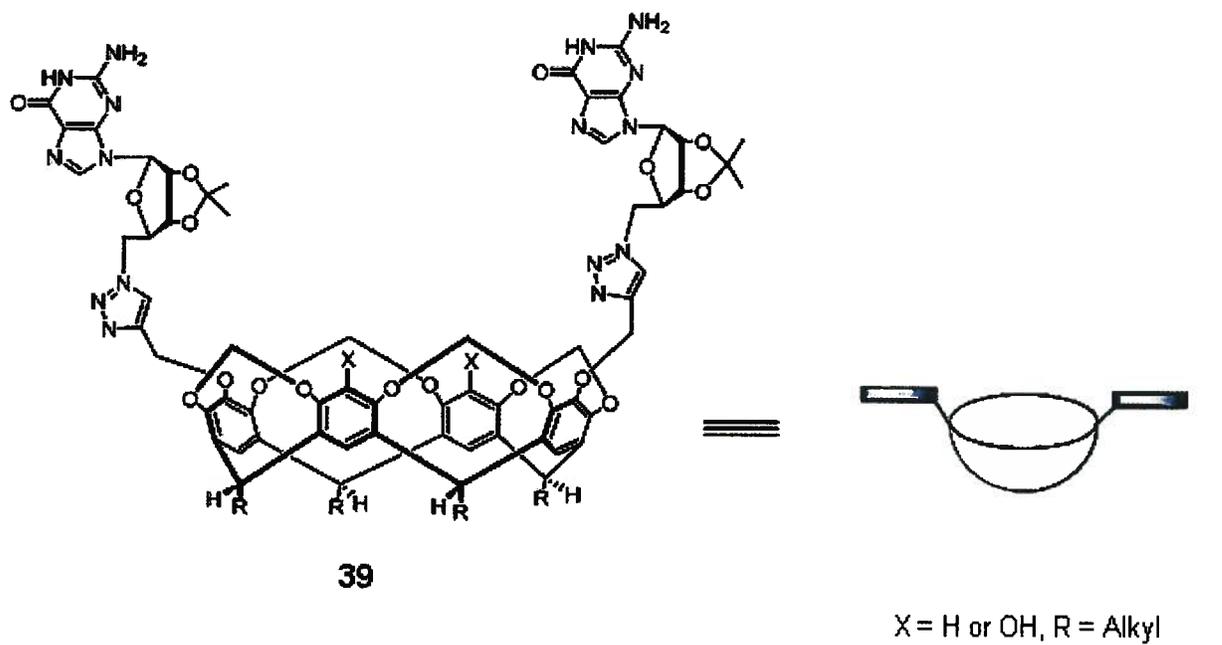
Although the lipophilic TASQs described in this thesis work have been characterized properly in the solution-phase, a crystal structure would still be an invaluable asset. Fortunately, TASQ·cation assemblies manifest a high potential for crystallization. More recently, single diffracting crystals of TASQ **10a**·Na<sup>+</sup> and TASQ **10a**·Cs<sup>+</sup> were grown in CHCl<sub>3</sub>/EtOH and work is in progress to process the collected datasets and to resolve the solid-phase structures.

### 5.2.1.2 Future Applications

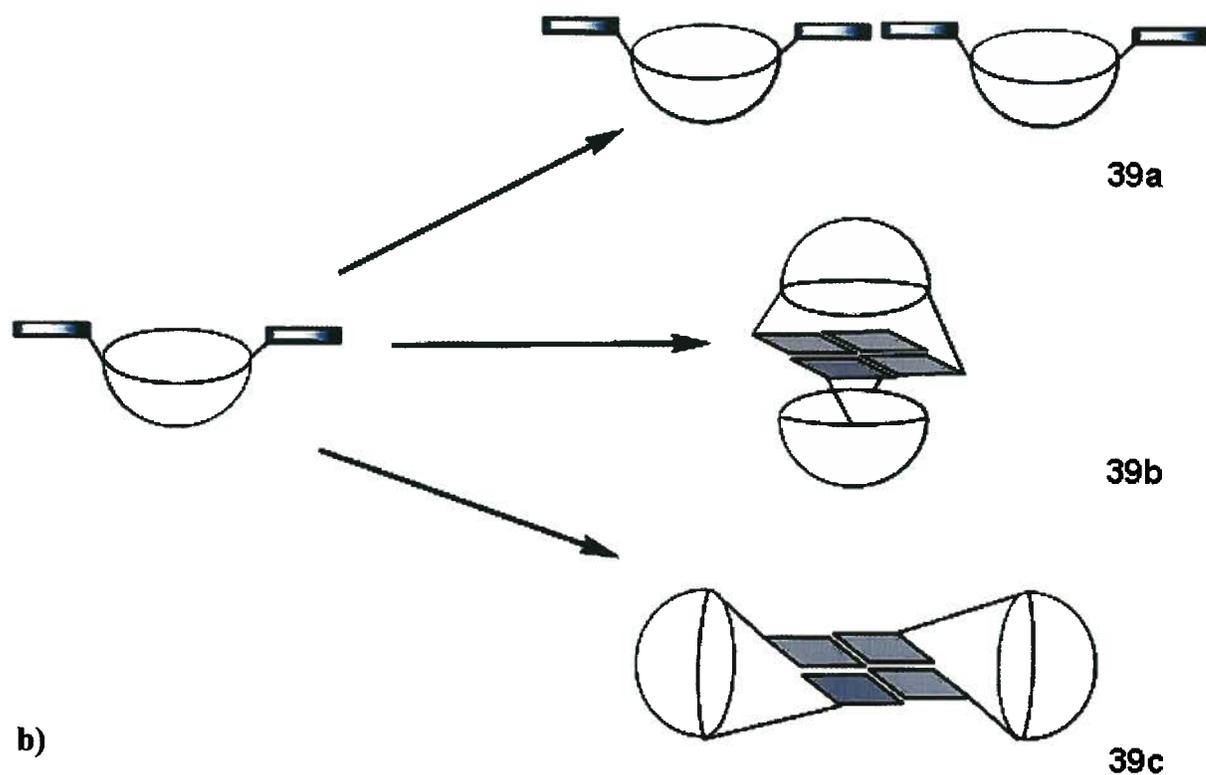
Work on lipophilic TASQs can be extended to different areas. For example, the lipophilic TASQs can be the basis of a cation-sensor or an ion-selective electrode. While the first generation of the lipophilic TASQs gave mainly monomeric and dimeric assemblies, they can be modified in a way to form a H-bonded polymer. Such cavity-bearing polymers, if formed, can in principle be used to adsorb or encapsulate the small guests.

Compound **39**, shown in Figure 5.1, can be readily synthesized based on the same chemistry described in this thesis and utilized for the formation of different architecture. The aggregation of **39**, depending on the conditions and the presence or absence of cations, may result in the formation of a linear polymer **39a**, a sandwiched dimer **39b**, or a stacked polymer **39c** (Figure 5.1). Although ditopic guanosine conjugates have been reported,<sup>11</sup> soluble polymeric systems such as **39a** and **39c** have not yet been prepared.

**Figure 5.1.** a) A ditopic TASQ monomer, b) Three possible arrays derived from self-assembly of compound **39**: a linear polymer **39a**, a sandwiched dimer **39b**, and a stacked polymer **39c** (only one stack is shown).



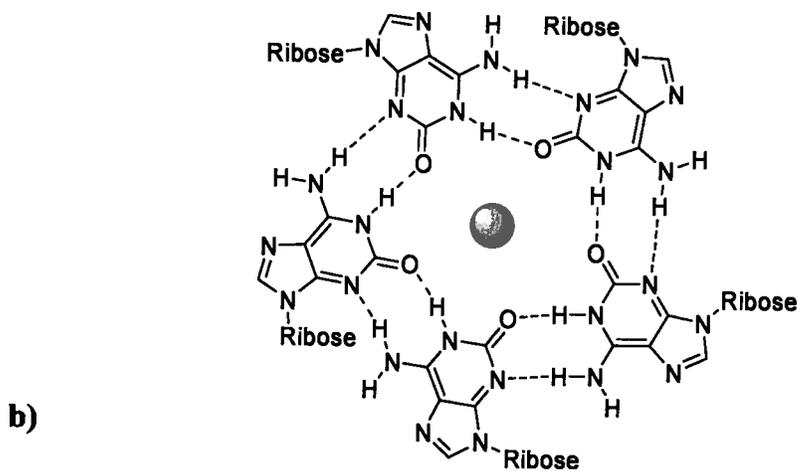
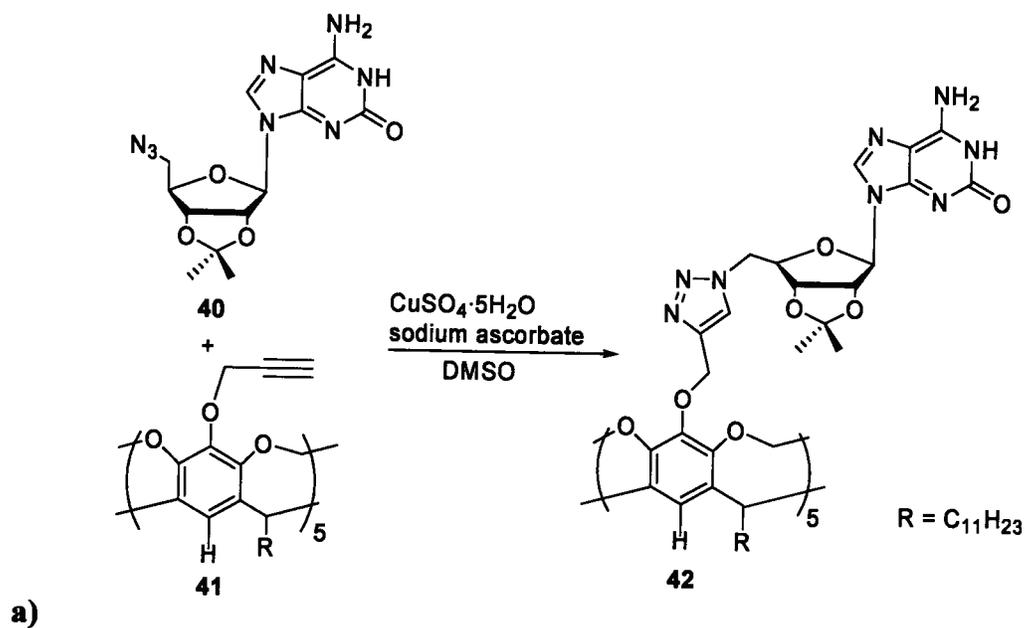
a)



b)

Another interesting area for future research would be the development of larger TASQs or TASQ analogs. The coupling of isoguanosine **40** to [5]cavitand **41** could afford a pentameric basket with adequate dimensions to be used as a microreactor or host cavity (Scheme 5.1). Given the established selectivity of isoguanosine for larger cations,<sup>12</sup> compound **42** could be a good candidate for the removal of  $^{137}\text{Cs}^+$  and  $^{226}\text{Ra}^{2+}$ .

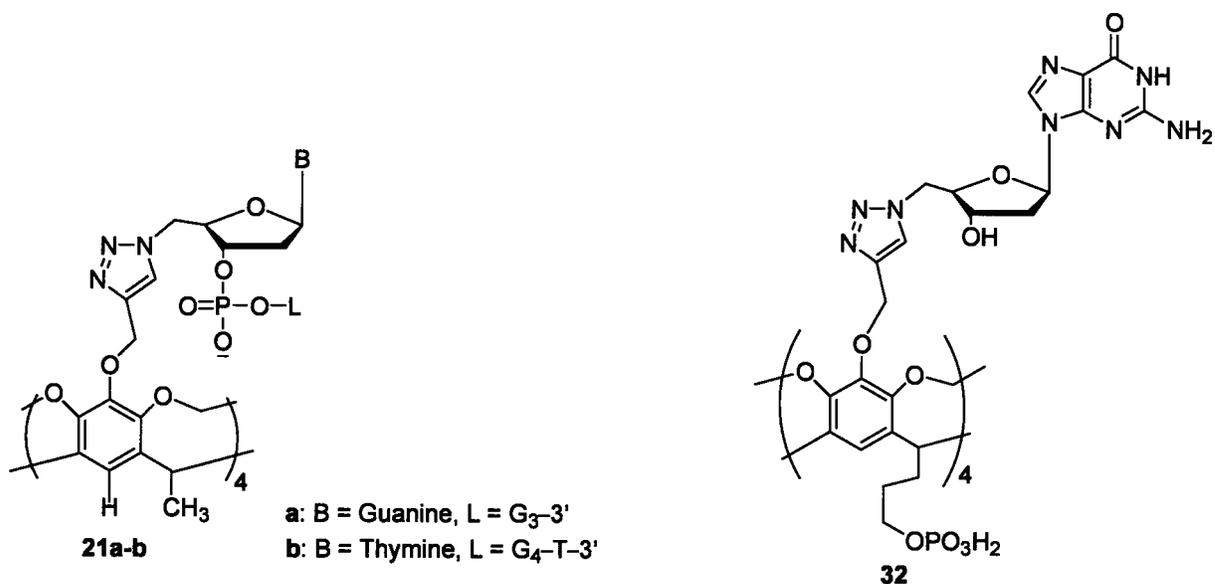
**Scheme 5.1.** a) Proposed synthesis of a pentameric TASQ analog using the “click” reaction, b) Isoguanosine pentamer.



## 5.2.2 Hydrophilic TASQs Project

### 5.2.2.1 Complementary Work

Immediate efforts should be directed towards synthesizing compounds **21a-b**, and **32** on Schemes 4.1, 4.3 respectively. For the synthesis of compounds **21a-b**, the influence of adjacent alkynes on coordination saturation of the copper catalyst needs to be studied. One quick way to examine this is to perform a reaction between a monopropargyl ether cavitand and a sample azide. The result of this experiment determines the next step which could be any of the following cases:



1) If the reaction is successful, then spatial proximity of the alkyne groups is most probably the problem preventing the reaction of the tetrasubstituted cavitand. Thus, switching the azide and alkyne functionalities of the reactants would be adequate to eliminate the problem. Various methods can be proposed to install the reactive groups based on the literature procedures

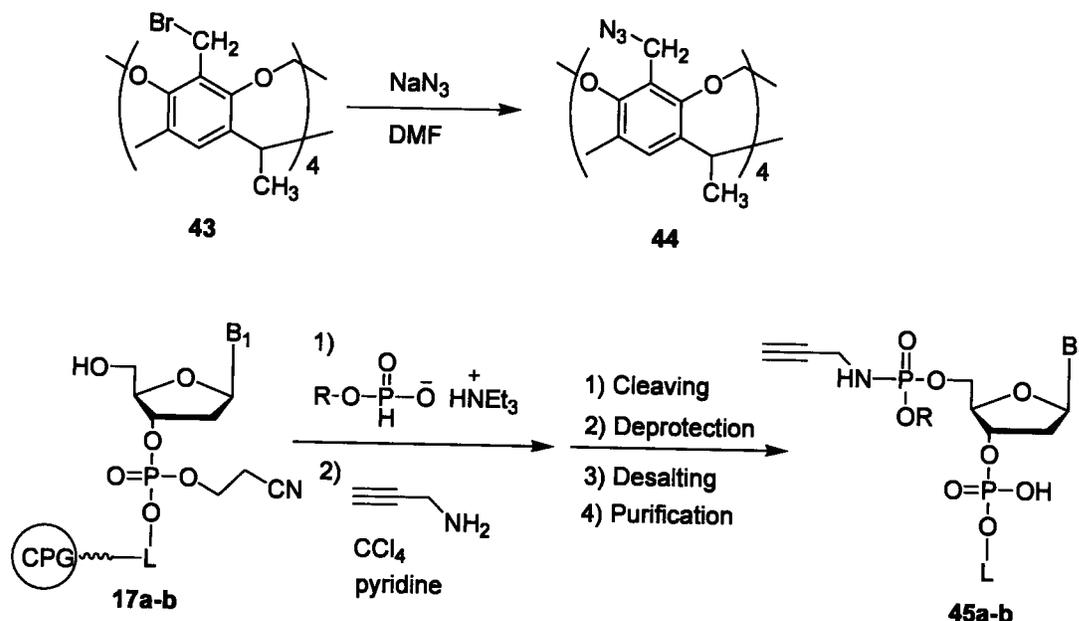
including those shown in Scheme 5.2. Compounds **43**<sup>13</sup> and **17a-b** are available and the synthetic methods proposed to give **44** and **45a-b**<sup>14</sup> are well established.

Assuming the validity of the above mentioned scenario and given the success of the “click” reaction in lipophilic media it would also be worthwhile to investigate the kinetic and mechanistic aspects of the “click” reaction of the cavitands. These studies can reveal the extent of the interaction of the copper catalyst with multiple alkyne centers under various conditions and eventually clarify a topic which has been a subject of contention in the literature.<sup>15</sup>

2) If the test reaction fails, then efforts must be focused on other areas. First, the ratios of the catalyst need to be explored as this is the important factor that determines the success of the reaction. In addition, the non-specific coordination of the copper ions to phosphates or guanines (although not clearly stated in the literature) need to be investigated. Beside Cu(I), the “click” reactions can also be promoted by some other catalysts. Metals such as Ru(II),<sup>16</sup> Pt(II) or Pd(II)<sup>17</sup> have been reported to catalyze the “click” reaction. Pt(II), for example, has shown promising results but produce slower reaction rates compared to those of Cu(I).<sup>17</sup> On the other hand, it has the advantage of eliminating the need for a reducing agent and exclusion of oxygen.

Apart from deactivation of the catalyst, the test reaction is expected to answer the concerns related to the steric accessibility of the reacting sites. This issue can also be further examined using cavitand with longer arms or more flexible DNA strands such as compound **20b** (Scheme 4.2). Finally, the reaction of the cavitand with the protected oligonucleotides **19a-b** (Scheme 4.2) can potentially be conducted on solid-support to facilitate the reaction and to provide easier separation of the products.

**Scheme 5.2.** Proposed reactants for the synthesis of a hydrophilic TASQ using the “click” reaction and their preparation methods.



**17a - 45a :** B<sub>1</sub> = Isobutyryl Guanine, B<sub>2</sub> = Guanine, L = G<sub>3</sub>-3', R = Alkyl  
**17b - 45b :** B<sub>1</sub> = Thymine, B<sub>2</sub> = Thymine, L = G<sub>4</sub>-T-3', R = Alkyl

Compound **32** (Scheme 4.3) shares some similarities with compounds **21a-b** (Scheme 4.1). These products are made from starting materials containing similar structural cores and identical reactive centers using “click” reaction. Thus, any synthetic or mechanistic information gained for one of the projects will have a substantial positive impact on the other project. From an analysis point of view, compound **32** offers some advantages as it can potentially be synthesized in larger scales (note that oligonucleotides are typically available in nM-mM scales). Thus, it would be easier to analyze the product and monitor the reaction by <sup>1</sup>H NMR, IR (disappearance of the azide band), or TLC. It would also be interesting to conduct a coupling reaction between the tetrapropargyl ether cavitand **30** and oligonucleotides **20a-b** and to compare the results to the coupling of the same scaffold to smaller azides.

### **5.2.2.2 Future Applications**

Following the synthesis of hydrophilic TASQs, a number of projects could be taken up including preparation of polymer-bound TASQs. Such systems may be served as stationary phases for the solid-phase separation of ligands and proteins from a library of compounds. TASQs may be used as vectors capable of interacting with DNA or may promote formation of long parallel polymers such as G-wires.

### 5.3 References

- (1) Neidle, S.; Balasubramanian, S. *Quadruplex nucleic acids*; RSC Pub.: Cambridge, 2006.
- (2) Arnal-Herault, C.; Pasc, A.; Michau, M.; Cot, D.; Petit, E.; Barboiu, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 8409-8413.
- (3) Alberti, P.; Bourdoncle, A.; Sacca, B.; Lacroix, L.; Mergny, J. L. *Org. Biomol. Chem.* **2006**, *4*, 3383-3391.
- (4) Pieraccini, S.; Gottarelli, G.; Mariani, P.; Masiero, S.; Saturni, L.; Spada, G. P. *Chirality* **2001**, *13*, 7-12.
- (5) Calzolari, A.; Di Felice, R.; Molinari, E.; Garbesi, A. *App. Phys. Lett.* **2002**, *80*, 3331-3333.
- (6) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.
- (7) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I.; Blackburn, E. H. *Cell* **1987**, *51*, 899-908.
- (8) Simonsson, T. *Biol. Chem.* **2001**, *382*, 621-628.
- (9) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596-2599.
- (10) Pinnavaia, T. J.; Marshall, C. L.; Mettler, C. M.; Fisk, C. I.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1978**, *100*, 3625-3627.
- (11) Kotch, F. W.; Sidorov, V.; Lam, Y. F.; Kayser, K. J.; Li, H.; Kaucher, M. S.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 15140-15150.
- (12) Cai, M.; Marlow, A. L.; Fettingner, J. C.; Fabris, D.; Haverlock, T. J.; Moyer, B. A.; Davis, J. T. *Angew. Chem., Int. Ed.* **2000**, *39*, 1283-+.
- (13) Sorrell, T. N.; Pigge, F. C. *J. Org. Chem.* **1993**, *58*, 784-785.
- (14) Bouillon, C.; Meyer, A.; Vidal, S.; Jochum, A.; Chevolut, Y.; Cloarec, J. P.; Praly, J. P.; Vasseur, J. J.; Morvan, F. *J. Org. Chem.* **2006**, *71*, 4700-4702.
- (15) Bock, V. D.; Perciaccante, R.; Jansen, T. P.; Hiemstra, H.; van Maarseveen, J. H. *Org. Lett.* **2006**, *8*, 919-922.
- (16) Zhang, L.; Chen, X. G.; Xue, P.; Sun, H. H. Y.; Williams, I. D.; Sharpless, K. B.; Fokin, V. V.; Jia, G. C. *J. Am. Chem. Soc.* **2005**, *127*, 15998-15999.

- (17) Golas, P. L.; Tsarevsky, N. V.; Sumerlin, B. S.; Matyjaszewski, K. *Macromolecules* **2006**, *39*, 6451-6457.