SYNTHESIS OF TEMPLATE-ASSEMBLED NUCLEOTIDE QUARTETS

by

Grant Bare

Bachelor of Science Honours, University of Saskatchewan, 2004

A THESIS SUBMITTED IN PARTIAL FULLFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Chemistry)

UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2012

Abstract

The thesis work described herein addresses a desire to isolate a guanine quartet and other unknown nucleobase quartets in Nature's solvent, water. Template-assembly of the nucleobase quartets or tetrads by the application of a cavitand template molecule was utilized as a synthetic strategy. Several synthetic precedents were sought and included developing DNA synthetic methodologies for four-fold nucleotide coupling of deoxyguanosine onto an individual substrate, imparting solubility in water on the resultant supramolecular constructs, and designing molecules capable of stabilizing G-quartet or G-tetrad self-assemblies.

Chapter 1 forms an introduction for this work by summarizing significant reports in the scientific literature on nucleic acid structure and function and on known template-assembled syntheses of nucleic acid elements. Chapter 2 describes the development of phosphite triester and phosphoramidite DNA synthetic methodologies for synthesizing water-soluble phosphate-linked cavitand-nucleotide conjugates. The synthesis of triazole-linked cavitand-guanosine conjugates with imparted water-solubility is dealt with in Chapter 3. In Chapter 4, evidence is provided for the observation of unimolecular cation-free template-assembled guanine and thymine quartets in methanol solvent using the phosphate-linked architecture. Finally, Chapter 5 describes the first potential isolation of a single G-tetrad in water through template-assembly directed by a phosphate-linked cavitand-nucleotide conjugate. This species is shown to interact with small molecule binders of G-quadruplex DNA and inhibitors of the human telomerase enzyme.

The novel supramolecules synthesized in this thesis have been shown to assist in the discovery of unusual DNA based structures, to act as a minimal model of human telomeric G-quadruplex DNA, and to serve as artificial receptors for potential anticancer therapeutics.

Preface

Parts of Chapter 3 were carried out in collaboration with Mehran Nikan and have been published elsewhere.[1] The synthesis of compounds **16** from **11** and **19b** from **16** was precedented by M. Nikan.[2,3] Compounds **11-16** and **19b** were also synthesized by the author. The published manuscript was written by the author, and all published ^{1}H NMR and HRMS characterization of compounds numbered here as **11-16** were performed by the author.[1] Characterization by ^{13}C NMR for compounds numbered here as **11-16** for publication and writing of the experimental procedures appearing in the supplementary data were carried out in collaboration with M. Nikan. Design, synthesis, and characterization by ^{1}H NMR, ^{13}C NMR, and MS of the published water-soluble triazole-linked cavitand-guanosine conjugate numbered here as **19a** was the sole work of the author. As well, the synthesis of compound **18** from **17** is novel to this author. All other work described in thesis chapters 2, 4, and 5 is original to the author and is unpublished.

(1) Nikan, M.; Bare, G. A. L.; Sherman, J. C. *Tetrahedron Lett.* **2011**, *52*, 1791-1793.

(2) Nikan, M. Dissertation, University of British Columbia, 2009.

(3) Nikan, M.; Sherman, J. C. unpublished results, 2010.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

1D	one-dimensional
2D	two-dimensional
A	adenine
Å	angstrom
Ac	acetyl
An	anisoyl
Bn	benzyl
br	broad
Bu	butyl
Bz	benzoyl
C	cytosine
CD	circular dichroism
CE	2-cyanoethyl
COSY	correlational spectroscopy
d	doublet
DCM	dichloromethane
DEAE	diethylaminoethyl
DIPEA	*N,N*-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethylsulfoxide

DMT	4,4'-dimethoxytrityl
DNA	deoxyribonucleic acid
DOSY	diffusion ordered nuclear magnetic resonance spectroscopy
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
Et	ethyl
GMP	guanosine monophosphate
h	hour
HRMS	high resolution mass spectrometry
i-Pr	isopropyl
Ibu	isobutyryl
IC_{50}	inhibition concentration required for 50 % of telomerase activity
m	multiplet
Me	methyl
min	minute
MALDI-TOF	matrix assisted laser desorption and ionization - time of flight
MMT	4-monomethoxytrityl
MS	mass spectrometry
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction

PDB	protein data bank
ROESY	rotating frame nuclear Overhauser effect spectroscopy
RNA	ribonucleic acid
s	singlet
rt	room temperature
t	triplet
T	thymine
t-Bu	*tert*-butyl
TASQ	template-assembled synthetic G-quartet
TBAF	tetra-*n*-butylammonium fluoride
TBS	*tert*-butyldimethylsilyl
TCE	2,2,2-trichloroethyl
TEA	triethylammonium
TEAB	triethylammonium bicarbonate
THF	tetrahydrofuran
TRAP	telomeric repeat amplification protocol
Tris	tris(hydroxymethyl)aminomethane
U	uracil
UV / VIS	ultraviolet / visible

Acknowledgments

I would like to foremost thank my supervisor Professor John Sherman for the inspiration for this thesis project and for providing an opportunity to carry out this research in his laboratory. I would like to thank past and current group members for helpful discussion and support including M. Nikan, J. Freeman, and B. Hui. The technical expertise provided by the UBC NMR and mass spectrometry laboratories staff, including Zorana Danilovic, Maria Ezhova, Yun Ling, Marshall Lapawa, and David Wong, has been an invaluable service. Zorana Danilovic designed ROESY NMR programs for the characterization of compound **3d**. Thank you to Brian Ditchburn for making many custom glassware pieces that made difficult synthetic steps manageable. Thanks goes to Bert Mueller of UBC Okanagan for performing all ICP-MS analysis. High resolution mass spectrometry of compound **19a** was carried out by the mass spectrometry staff in the Department of Chemistry at the University of Alberta. Finally, I am grateful for financial support that NSERC has provided to me through several scholarships.

Chapter 1: Introduction

1.1 An Introduction to DNA structure: Unraveling the B-DNA Helix

Perhaps a no more fitting example exists than that of the structure of DNA for the observation that in Nature simplicity begets complexity. In a pair of now celebrated papers put forth by J.D. Watson and F.H.C. Crick in 1953 on how the polynucleotide strands of the salt of DNA could associate, the universal basis of molecular genetics for all known living organisms was revealed.[1,2] At that time, the chemical constitution of DNA had already been recognized correctly as a polymer of 2'-deoxyribonucleosides connected through 3',5' phosphodiester linkages (Figure 1.1). The 2'-deoxyribonucleosides are a class of molecules composed from a D-ribose sugar attached at the C1' atom to the N1 or N9 atom of a respective pyrimidine or purine nitrogenous base in a β- configuration. In their model of what is now referred to as B-DNA or physiological DNA, Watson and Crick proposed that two antiparallel polynucleotide strands could associate non-covalently to form an asymmetric right-handed double helical structure with the nitrogenous bases lying perpendicular to the helical axis (Figure 1.2). The nitrogenous bases associated in one of two ways through hydrogen bonded complementary base paring. Only thymine could go with adenine, and only cytosine could go with guanine. The resulting helix repeated itself every 10 residues or 34 Å. Several observations were cited in the deduction of this structure. Chargaff had reported

Figure 1.1 a) The purine and pyrimidine nitrogeneous bases of DNA. b) Examples of the 2'-deoxyribonucleosides: 2'-deoxycytidine and 2'-deoxyadenosine. c) A single polynucleotide strand of DNA depicted in the 5'-3' sense.
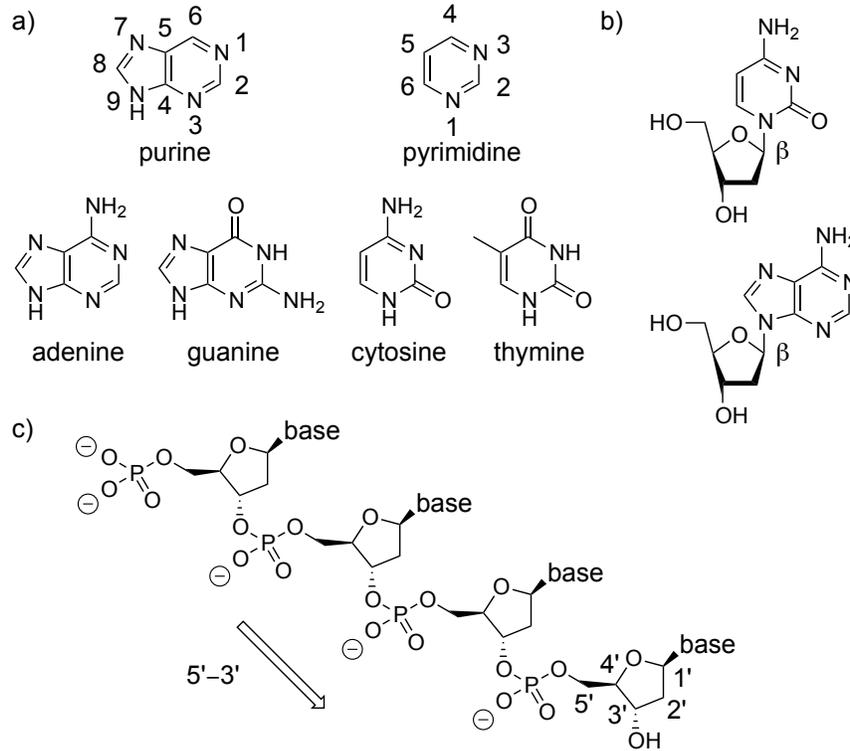
that the molar ratio of thymine to adenine and cytosine to guanine in microbial DNA was close to unity.[3] Guanine and thymine were assumed to exist exclusively in their more stable keto tautomeric forms. Finally, data from colleagues at Cambridge concerning the X-ray diffraction of DNA fibers was consistent with a two stranded helix.[4,5]

The solution to two mysteries of molecular genetics became immediately apparent from the proposal of the B-DNA model. Firstly, the belief that the sequence of bases was irregular suggested to them that the genetic information of an organism lies in its

2

Figure 1.2   a) The double helical structure of B-form DNA proposed by Watson and Crick. b)

The **A · T** and **G · C** complementary Watson-Crick base pairs.
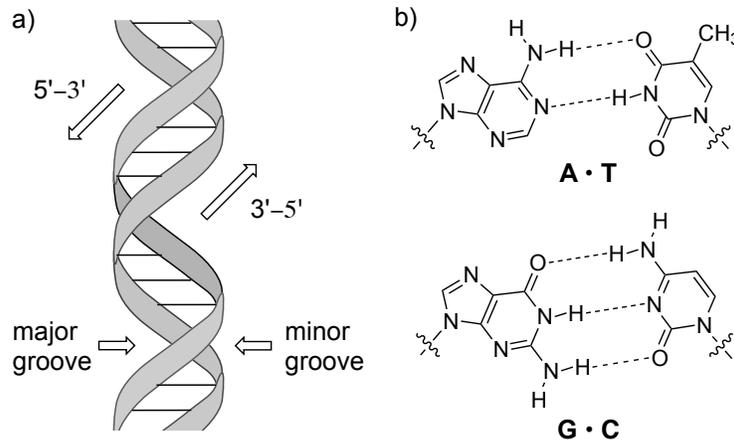
sequence of bases. Secondly, the complementary nature of the two strands of the DNA molecule provided a mechanism whereby each strand could in turn act as a template for its complementary strand allowing the exact duplication of the DNA molecule or the genetic information found within it. Thus, the simple model of DNA could now account for the tremendous diversity observed among known living organisms.

A structural proof of the B-DNA helix would prove to be elusive due in part to a lag in the development of DNA synthetic methodologies that were suitable for producing sufficient amounts of pure sequence for crystallization and X-ray analysis. High resolution X-ray crystallographic structures of 1-methylthymine and 9-methyladenine confirming the existence of a Watson-Crick base pair and the existence of an alternative pairing type referred to as Hoogsteen base pairing that involved hydrogen bonding at the non Watson-Crick face were known soon after Watson and Crick's proposal (Figure

3

1.3).[6,7] However, a confirmation of the structure of Watson and Crick's B-DNA was not achieved until 28 years after its discovery.

Figure 1.3   A Hoogsteen type base pair between 1-methylthymine and 9-methyladenine.
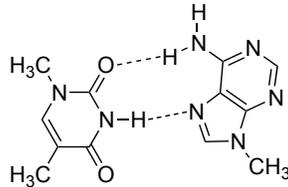
The first single crystal structure of B-DNA was of a self-complementary dodecamer of sequence d(CGCGAATTCGCG) that incorporated the EcoRI restriction endonuclease site, GAATTC (Figure 1.4).[8,9] This structure was found to be similar to the idealized B-form structure in many respects. It was a right-handed double helix with 10.1 base pairs per turn with a height of 3.4 Å per base pair. The *N*-glycosidic torsion angles were orientated *anti* as opposed to *syn*. Sugar pucker conformations were distributed closer to C2'-endo and not the C3'-endo conformation found in an A-form DNA helix. A departure from the classical B-DNA model was evident however from the observation that the Watson-Crick base pairs deviated from planarity and propeller-twisted in the crystal structure.

Figure 1.4 X-ray crystallographic structure of the B-DNA dodecamer d(CGCGAATTCGCG) (PDB entry 1BNA) looking a) towards helical axis and b) down helical axis modeled with UCSF Chimera, a visualization system for exploratory research and analysis. (Pettersen E.F.; Goddard T.D.; Huang C.C.; Couch G.S.; Greenblatt D.M.; Meng E.C.; Ferrin T.E. *J. Comput. Chem.* **2004**, *25*, 1605-1612). Important ribose sugar conformations in DNA and RNA structure: c) *N*-glycosidic torsion angles and d) sugar ring puckers.

a)

b)

c)

anti

anti

syn

syn

d)

2'
5'
4'
1'
3'
base
C2'-endo

3'
5'
4'
2'
base
C3'-endo
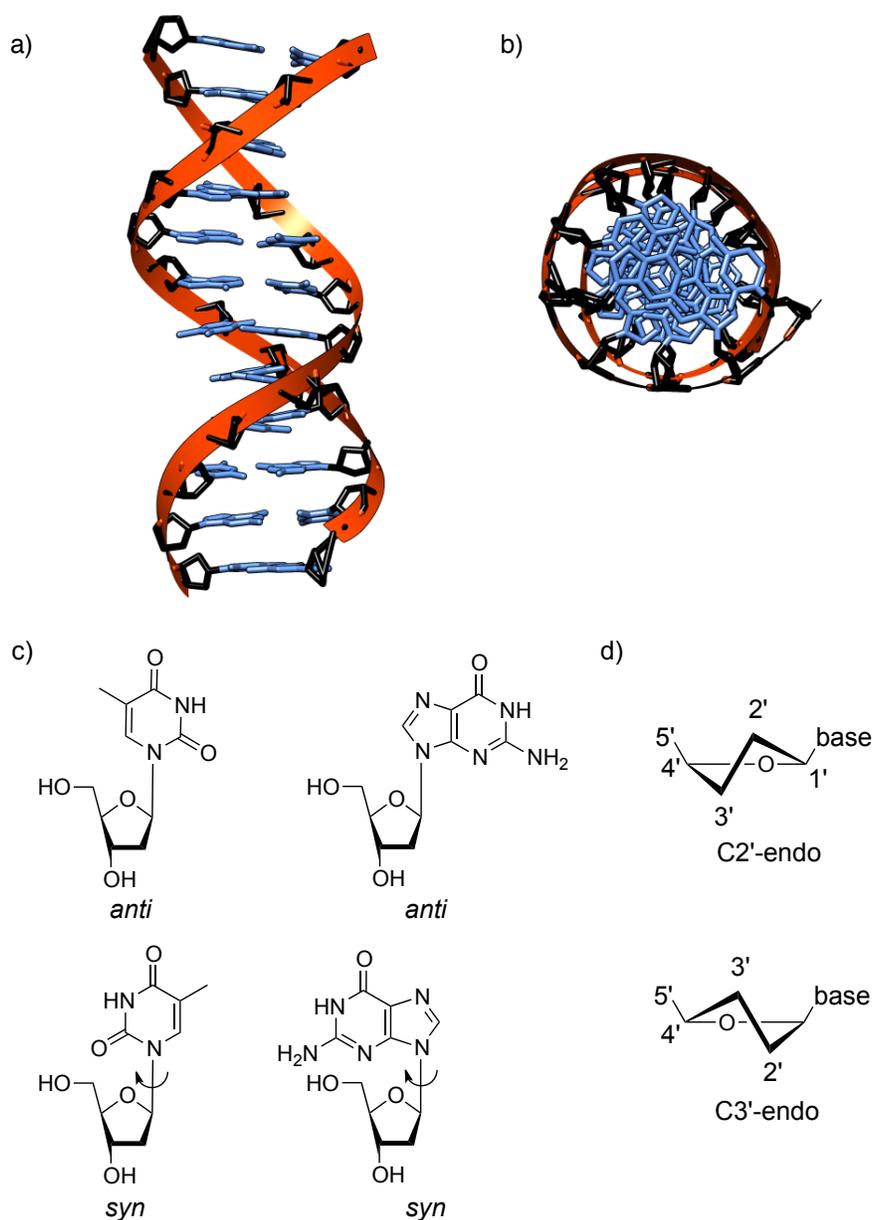
1.2   A New Paradigm in DNA Structure: The Guanine Tetramer

It was known as early as 1910 that concentrated solutions of guanylic acid exhibited the peculiar property of gel formation.[10] In 1962, Ralph and coworkers observed that DNA oligonucleotides containing three and four guanines and with a phosphate at the 5' position showed unusually stable secondary structure that was absent in the analogous oligonucleotides containing thymine, adenine, or cytosine.[11] At the same time, Gellert reported that concentrated solutions of guanylic acid, specifically 5'-GMP and 3'-GMP, formed gels at pH 5 in the presence of sodium chloride salt.[12] Based primarily on X-ray diffraction data and modeling studies inspired by Donohue's comprehensive investigation of base pairing in DNA polynucleotides, Gellert proposed that the remarkable properties of guanine shown thus far could be explained by the formation of hydrogen bonded planar guanine tetramers with four fold rotational symmetry (Figure 1.5).[13] In this arrangement, the 2-amino hydrogen could donate one hydrogen bond to the 7-nitrogen of a neighboring residue, and the 1-imino hydrogen could donate one hydrogen bond to the 6-carbonyl oxygen of the same residue. Guanine is the only base to possess two sets of hydrogen bond donor-acceptor pairs oriented at near right angles with respect to each other. These planar guanine tetramers can stack favorably, and in the case of 5'-GMP, Gellert pointed out that the existence of a left-handed helix where each tetramer has a helical rotation of 22.5° and an axial translation of 3.25 Å would be consistent with the X-ray diffraction results.

The first direct evidence that polynucleotide strands could indeed assemble into helical structures of stacked guanine tetramers came from studies of RNA

Figure 1.5   a) 5' and 3' guanosine monophosphates observed in gel formation. Structures

proposed by Gellert: b) G-tetrad and c) helical stack of G-tetrads.

a)

b)

c)
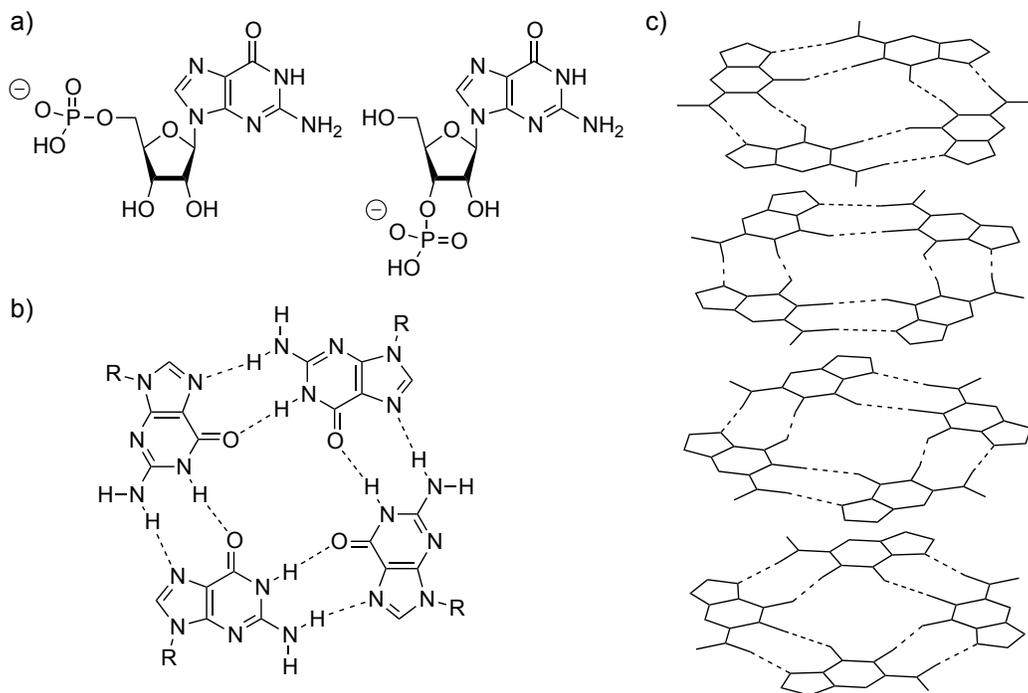
polynucleotides. A four-stranded model was suggested to explain the X-ray diffraction

patterns obtained from salts of polyguanylic acid (Figure 1.6).[14] In this model, four

individual polynucleotide strands can be envisioned to associate through four in-plane

hydrogen bonded guanine residues where each residue is donated by one of the four

polynucleotide strands. A helix structure results with a helical rotation of 31.2° per

tetramer. Certain stereochemical requirements needed to be satisfied in the proposed

structure. Guanine residues needed to be oriented *anti* with respect to the glycosidic bond

conformation, and the steric bulk of the 2' alcohol required that sugars adopt a C3'-endo

conformation. Helical stacks of guanine tetrads either covalently linked as in the

7

Figure 1.6   a) Various G-quadruplex strand topologies: a) four-stranded all parallel as in

Zimmerman's poly r(G) or Sen's immunoglobin switch region, b) intramolecular

antiparallel reported by Williamson for DNA strand d[(T$_4$G$_4$)$_4$], and c)

antiparallel dimer reported by Sundquist for DNA strand d[(T$_2$G$_4$T$_2$G$_4$)$_2$]. Helical

twist is not shown for simplicity. d) Alternating *N*-glycosidic conformations within

G-quartet plane of an antiparallel structure. Sodium cation is bound in-plane.

e) Coordination of potassium cation between G-quartet planes.
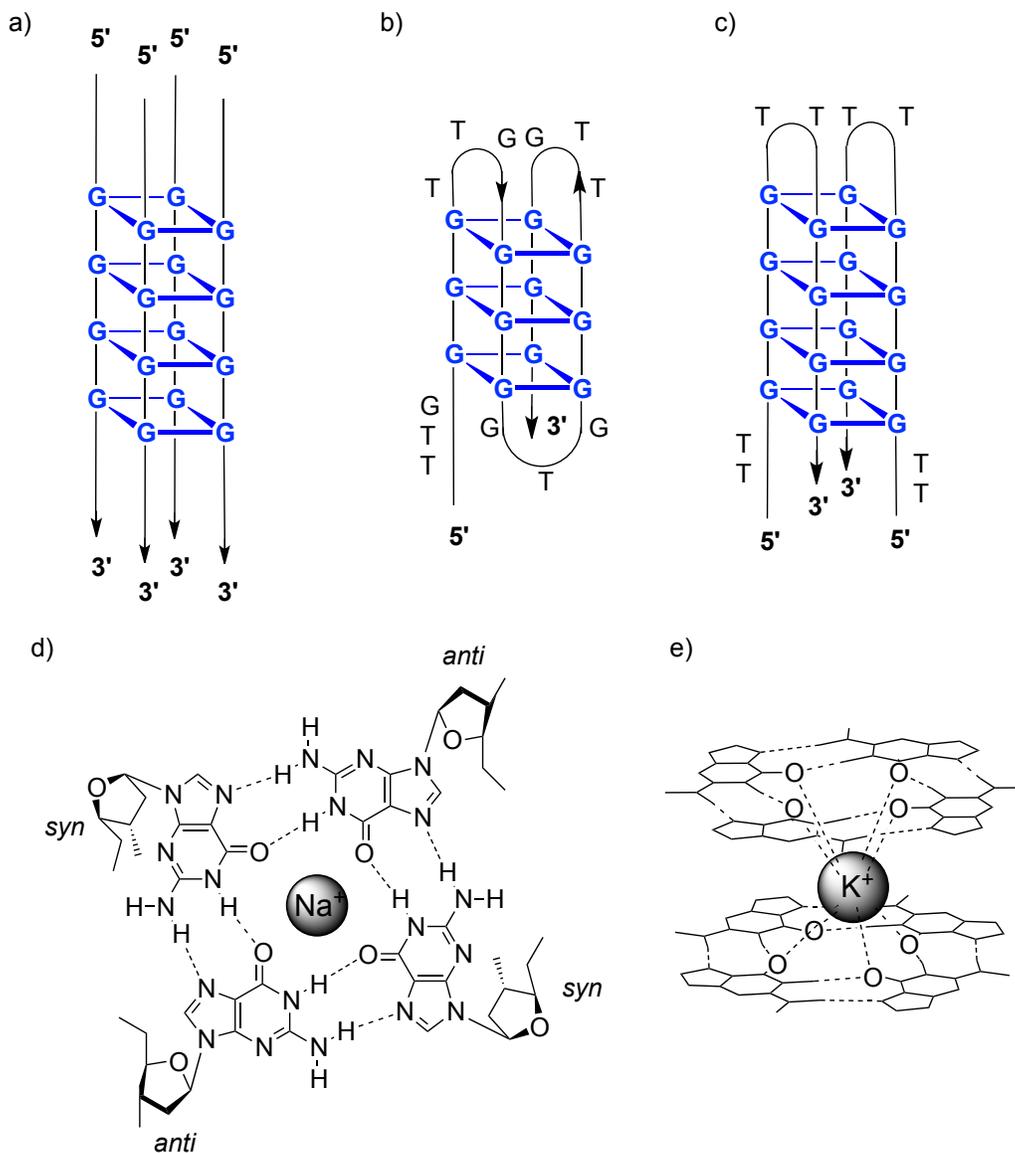
case of Zimmerman's poly r(G) oligomers or non-covalently linked like Gellert's ribonucleosides are commonly referred to as G-quadruplexes.

The investigation by Henderson *et al.* in 1987 on the physical properties of synthetic DNA oligonucleotides containing the guanine rich telomeric sequences of several organisms, including the ciliated protozoans *Oxytricha* and *Tetrahymena*, by gel electrophoresis and NMR methods revealed guanine's potential *in vivo* significance.[15] Although they did not go as far as to conclude the existence of guanine tetramers in these oligonucleotides, they did however provide evidence for the stabilization of intramolecular structures by non Watson-Crick guanine-guanine hydrogen bonded base pairs in the *syn* conformation. Shortly thereafter, Sen and Gilbert reported gel electrophoresis and N7 methylation experiments on single stranded DNA containing a guanine rich repeating motif found in the immunoglobulin switch region.[16] It was concluded that under physiological salt concentrations, these strands were forming parallel four-stranded structures containing adjacent hydrogen bonded guanine tetrads resembling the model presented by Zimmerman for polyguanylic acid.

A further investigation of telomeric sequences found in *Oxytricha* and *Tetrahymena* by Williamson and coworkers showed that individual DNA oligonucleotides containing several repeats of these sequences could be induced in the presence of sodium, potassium, or cesium cations, but not lithium cations, to self-fold into a new type of antiparallel structure.[17] In this antiparallel helix, strands alternate in sense and with respect to glycosidic torsion angle, such that in any guanine tetramer or "G-quartet" as they were coined, the residues will alternate between *syn* and *anti*. A binding site for the monovalent cation was proposed to be the G-quartet itself. The cation

selectivity was rationalized by the fact that lithium would be too small to fit in the center of the quartet, while sodium fit in plane and potassium and cesium fit out of plane. The observation that *Tetrahymena* telomeric DNA could also dimerize to form an antiparallel structure was reported by Sundquist.[18] Sundquist further refined the cation binding model. For example, potassium was believed to sit between planes of G-tetrads such that it was coordinated to all eight carbonyl oxygens in an octahedral cage.

1.3　The "End-Replication Problem", Telomeres, Telomerase, and the Connection to G-Quartet DNA

It was Watson, while studying the replication of linear viral T7 DNA together with the Russian biochemist, Olovnikov, who pointed out what is known to molecular biologists as the "end-replication problem" (Figure 1.7).[19-21] During DNA replication, daughter strand synthesis by DNA polymerases was known to proceed in the 5'-3' direction by adding incoming nucleotides to an already existing 3'-OH on the daughter strand polynucleotide chain. The requirement for a 3'-OH substrate by the DNA polymerase was satisfied by the placement of RNA primers, for example those found in Okazaki's pieces, during the initiation of strand synthesis.[22] However, it was realized that the removal of the terminal 5' RNA primer on either daughter strand by subsequent enzymatic degradation would give rise to a 5' gap and a 3' overhanging single strand. Thus, each round of replication of DNA would result in a progressively shorter

Figure 1.7   The end-replication problem: DNA replication of linear chromosome ends leads

to a 3' overhang on the parent strand and a 5' gap on the daughter strand.

duplex parent DNA

3' linear end
5'

daughter strand synthesis
initiation by RNA primers

3' lagging strand Okazaki pieces
5'

3' leading strand
5'

strand elongation by DNA polymerases

3'
5'

3'
5'

completion of strand synthesis,
removal of RNA primers

3' 3' overhang / 5' gap
5'

3'
5'
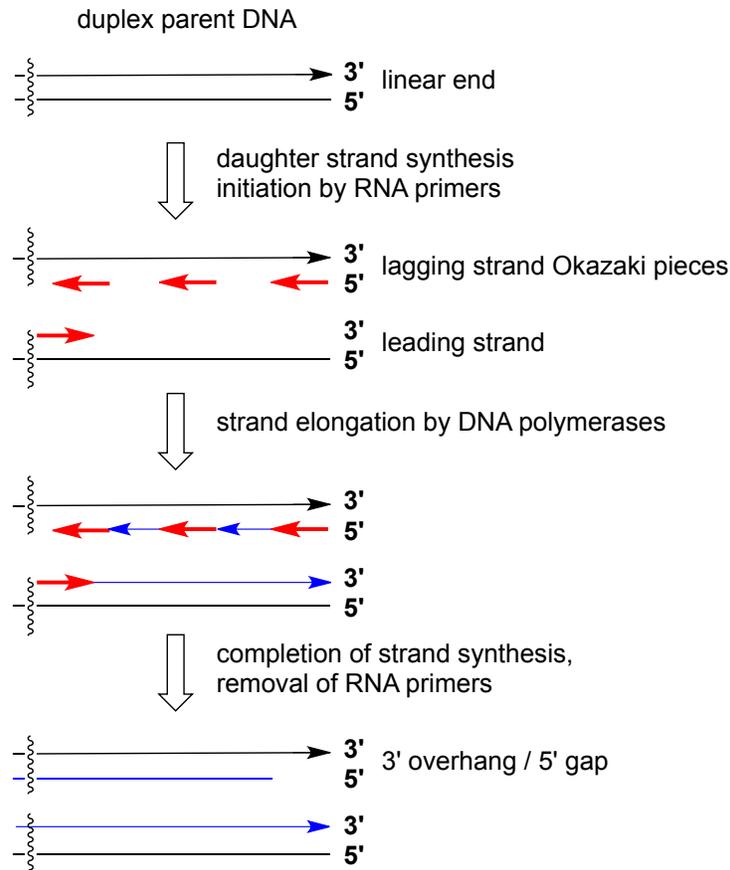
chromosome end. Organisms containing circular DNA genomes such as those found in

bacteria and in some viruses would be exempt from this end degradation process during

DNA replication.

Elucidation of the structure and maintenance of the end or telomeric region of

linear chromosomes was necessary to provide a solution to the end-replication problem.

Telomeres can be defined as the DNA elements belonging to the end region of a linear

11

chromosome that impart stability on the chromosome together with the proteins that are necessary to maintain these DNA elements.[23, 24] Telomeres perform a protective function against chromosomal end degradation and end-end chromosomal fusion events, whereas in contrast, double-stranded breaks in chromosomes lead to instability.[25]

In 1978, Blackburn and Gall identified a hexanucleotide sequence, CCCCAA, in the protozoan *Tetrahymena* that was tandemly repeated 20 to 70 times at the termini of linear ribosomal DNA.[26] Szostak and Blackburn shortly thereafter reported the construction of a stable linear yeast plasmid after attaching a fragment containing the *Tetrahymena* terminal ribosomal DNA hexanucleotide repeat to the ends of a yeast chromosome.[27] Astonishing was the observation that the yeast cellular machinery was able to recognize and elongate the *Tetrahymena* sequence with a distinct yeast telomeric sequence. Thus, the *Tetrahymena* hexanucleotide repeat not only represented a stable telomere, it also suggested that telomeres were highly conserved in evolution between two distantly related organisms like a protozoan and yeast.

The discovery of telomere terminal transferase or telomerase activity in *Tetrahymena* extracts helped to provide further insight into the mechanism of telomere elongation.[28] A telomere elongating enzyme that could catalyze the addition of the *Tetrahymena* telomere repeat TTGGGG of the G-rich strand onto an appropriate substrate had now been identified. The fact that only synthetic single stranded DNA primers containing the G-rich strand tandem repeats and not the C-rich repeats of the *Tetrahymena* and yeast telomeres, and that only dGTP and dTTP could act as substrates for the *Tetrahymena* extracts demonstrated that the G-rich strand was the functional telomere strand. The G-rich strand of the telomere has always been found to be oriented

12

in the 5'-3' sense toward the end of the chromosome and to form a 3' overhang beyond the C-rich strand for a wide range of eukaryotic organisms.[24] *Tetrahymena* telomerase was identified as a ribonucleoprotein complex[29] with an internal RNA template[30] of sequence CAACCCAA complementary to the G-rich telomere sequence essential for activity. Furthermore, the identification of an internal template suggested a mechanism for telomerase function where single-stranded telomere elongation could proceed by alternating rounds of reverse transcriptase-catalyzed elongation in the standard 5'-3' direction and translocation of the telomerase enzyme on the 3' overhang telomere substrate (Figure 1.8).

The human telomere is found to contain approximately five to fifteen kilobases of the repeating sequence TTAGGG, a sequence that is conserved in vertebrates, and to consist of atypically long 3' overhangs of the G-rich strand.[31,32] For example, Makarov reported 3' overhangs with average lengths of approximately 100 to 200 nucleotides in human telomeres.[33] Human telomerase like the telomerase identified beforehand in *Tetrahymena* is a ribonucleoprotein complex with a complementary internal RNA template.[17,34] The enzyme uses its 11 nucleotide template with sequence CUAACCCUAAC to elongate the G-rich 3' overhang substrate by successively adding TTAGGG units through its reverse transcriptase and translocase activities. Many proteins are found to be associated with human telomeres in addition to telomerase, and some of these proteins are necessary for proper function of human telomeres. Among them is the widely studied shelterin complex, a six subunit telomere binding protein that imparts stability on the telomere.[35,36]

Figure 1.8   The maintenance of *Tetrahymena* telomeres by telomerase enzyme requires

reverse transcriptase elongation and translocation activities.

telomerase enzyme

tetrahymena telomere

CCCCAACCCCAACCC **5'**   **5' AACCCCAAC 3'**
GGGGTTGGGGTTGGGGTTGGGGTTGGGG **3'**

⬇ elongation

CCCCAACCCCAACCC   **AACCCCAAC**
GGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG

⬇ translocation

CCCCAACCCCAACCC   **AACCCCAAC**
GGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG

⬇ elongation

CCCCAACCCCAACCC   **AACCCCAAC**
GGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG
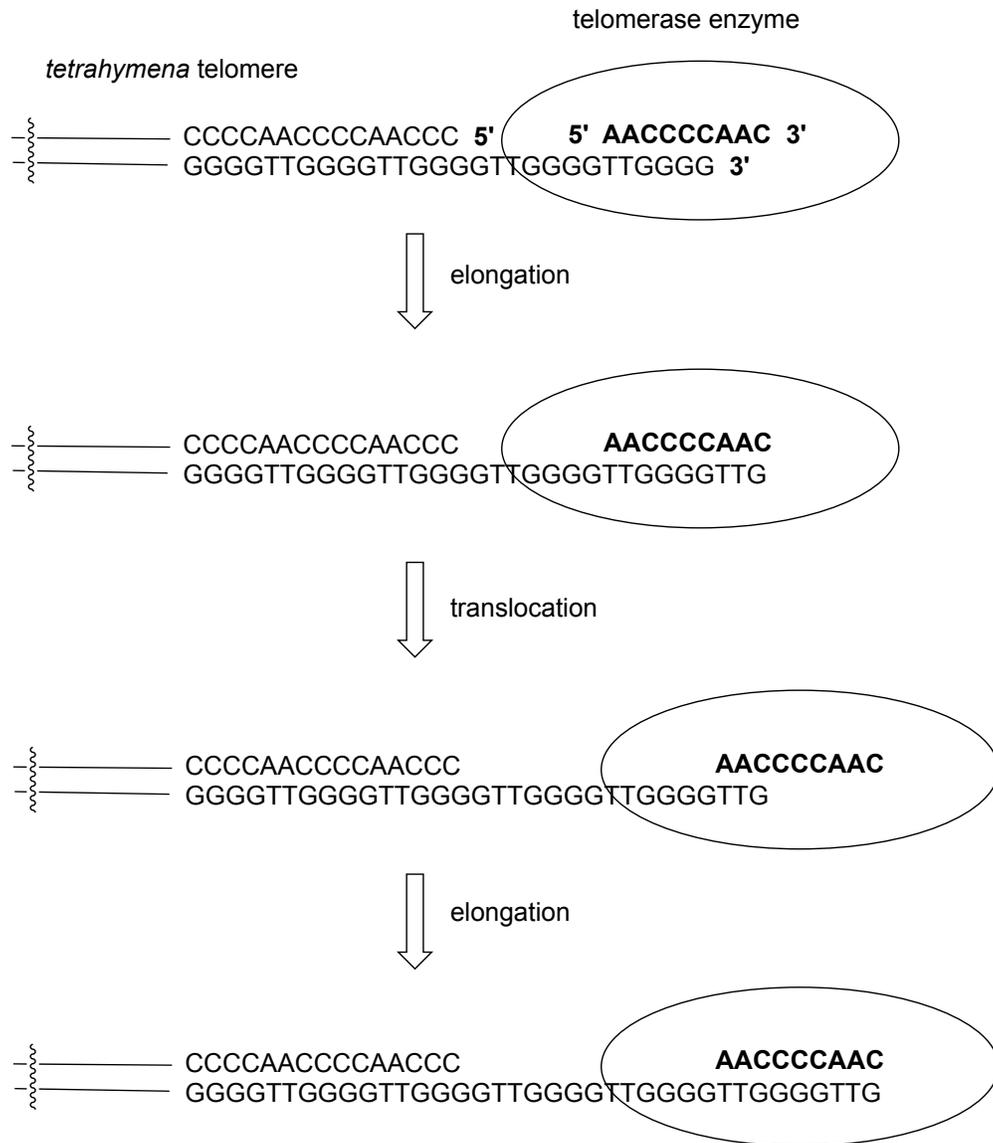
The existence of the repeating guanine run sequence motif in the 3' overhanging

strand of all studied telomeres is not accidental. Synthetic DNA oligonucleotides

comprising the telomere repeating sequences of ciliated protozoans and of humans have

been demonstrated *in vitro* to form G-quadruplex structures in the solid-state through high-resolution X-ray crystallography and in the solution-state by NMR methods (Figure 1.9). The first high-resolution X-ray crystal structure of a G-quadruplex was that published by Kang and coworkers for the *Oxytricha* telomere oligonucleotide d(GGGGTTTTGGGG).[37] The four-stranded antiparallel quadruplex structure was formed by dimerization of two hairpin-forming thymine-looped oligonucleotides. Glycosidic bond conformations were found expectedly to alternate between *syn* and *anti* within the guanine quartet. However, novel to all proposed models of G-quadruplex DNA, was the presence of alternating *syn / anti* conformations within the chains. Evidence for coordination of a potassium cation in-center but off-plane relative to the guanine tetrad was also presented. A short oligonucleotide containing the *Tetrahymena* telomere repeat, d(TGGGGT), also was crystallized into G-quadruplex DNA.[38] The resulting right-handed tetrameric structure was found to contain all parallel strands and *anti* glycosidic torsion angles. All but one sodium cation was coordinated with the eight carbonyls between two planes of guanine quartets.

The earliest report of a human telomere structure was by Wang *et al.* for solving the solution structure of d[AGGG(TTAGGG)₃] by NMR and molecular dynamics.[39] In the presence of sodium, an intramolecular G-quadruplex with three stacked G-tetrads and connected through TTA loops was elucidated. The right-handed twisted helix revealed asymmetrical grooves, parallel / antiparallel / antiparallel / parallel folding topology, and C2'-endo sugars with *syn / anti / anti / syn* conformations. A decade later, the same oligonucleotide in addition to a shortened variant with sequence d(TAGGGTTAGGGT)

Figure 1.9   Solid-state and solutions structures of DNA oligonucleotides containing telomeric

sequences: a) *Oxytricha* d[(G$_4$T$_4$G$_4$)$_2$] (PDB entry 1D59) b) *Tetrahymena*

d[(TG$_4$T)$_4$] (PDB entry 244D) c) human d[AG$_3$(TTAG$_3$)$_3$] (PDB entry 143D) d)

human d[(TAG$_3$TTAG$_3$T)$_2$] (PDB entry 1K8P) modeled with UCSF Chimera.
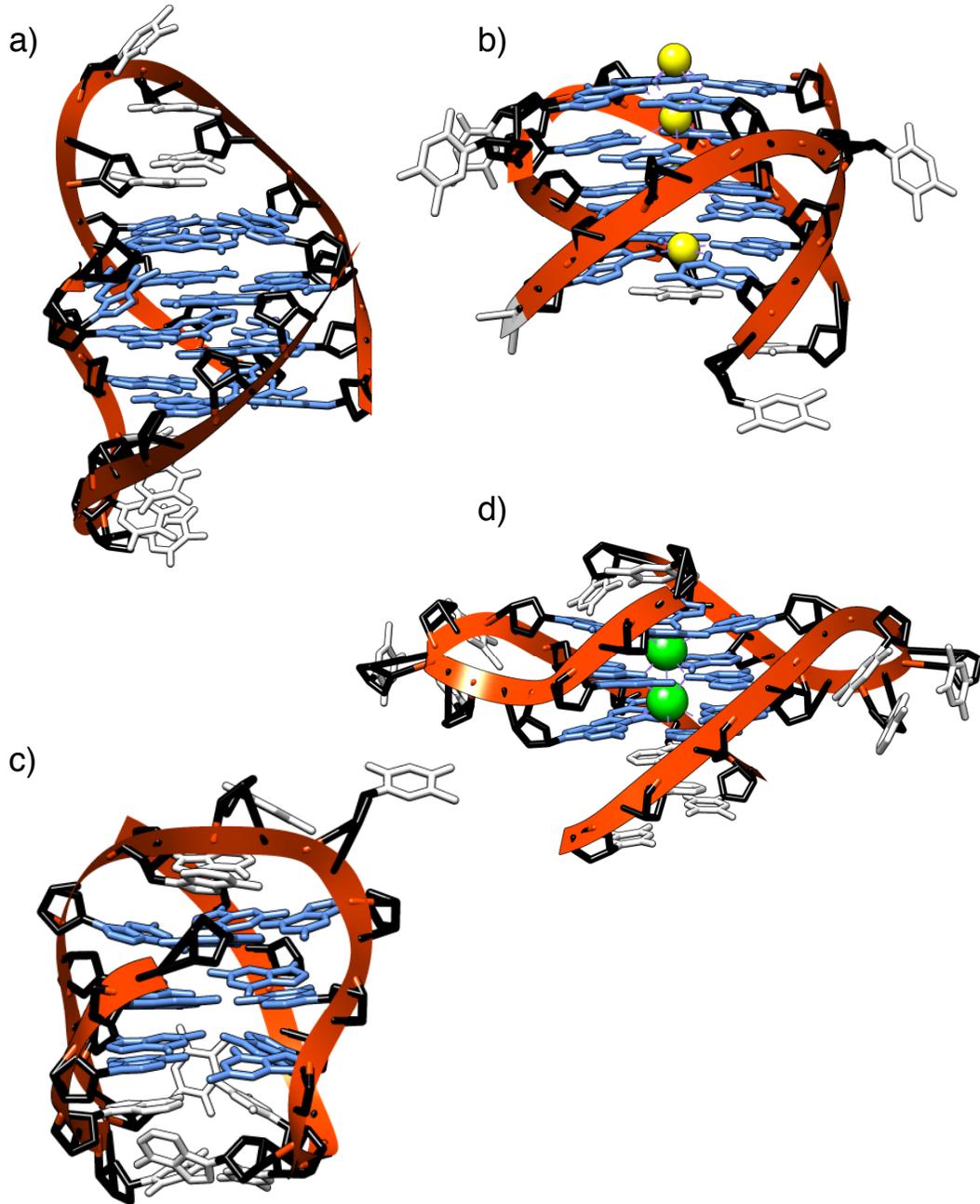
were crystallized.[40] Consistent with the solution structure beforehand was an intermolecular twisted three-layered quadruplex with C2'-endo sugar puckers. The shortened strand made an intermolecular dimer, but shared the same structural features of the longer strand. Dramatic differences in folding topology were observed however in the solid-state. An all parallel all *anti* structure was apparent with unique propeller loops connecting adjacent chains. Potassium cations were coordinated to eight carbonyls in the predicted bipyramidal antiprismatic fashion.

Direct *in vivo* evidence for G-quadruplex DNA in the genome of eukaryotic organisms has proved to be all but elusive. Representative work is given from the study of the interaction of G-quadruplex specific antibodies with DNA inside the cell nucleus. Immunofluorescence was observed when *in vitro* generated antibodies specific to G-quadruplex formed from the telomere of hypotrichous ciliated protozoan, d(TTTTGGGG), were reacted with the macronuclei of *Stylonychia lemnae*.[41]

Sequences of DNA with G-quadruplex forming potential are not unique to the telomere regions in genomic DNA. Huppert has proposed that G-quadruplex-forming motifs are prevalent in the human genome outside the telomeres and are localized to over 40 % of gene promoter regions.[42] Moreover, promoter G-quadruplex structures are believed to be involved with the control of gene transcription. The promoter of the human oncogene c-myc is the most widely studied of these types of promoter regions. Formation of G-quadruplex by c-myc promoter elements *in vitro*,[43] and the transcriptional suppression and activation of the c-myc gene by respective G-quadruplex stabilizing and destabilizing conditions has been reported.[44, 45]

1.4 Targeting Telomeres and Telomerase as a Novel Strategy for the Development of Cancer Therapeutics

It has been known for over a half-century that normal human cells have a finite capacity for dividing under laboratory conditions. Pioneering work undertaken by Hayflick demonstrated that many strains of human fibroblasts degenerate after approximately 50 cultivations or one-year of storage, and that this degeneration or cellular senescence as it was called has an intrinsic cellular origin.[46] The connection between cellular aging and cancer processes to telomere maintenance was not made until several decades following Hayflick's analysis. In 1990, Harley showed that telomere shortening could be correlated to the age of human fibroblasts.[47] Soon after, certain cells were observed to become immortalized following their entry into a crisis state after their telomeres had become sufficiently shortened due to excessive doublings.[48] The overexpression of telomerase activity was found in these immortalized cells, whereas no significant telomerase activity was found in control cells.

A rigorous investigation of telomerase activity over a wide range of human cell types by Kim and coworkers has produced the most comprehensive results linking telomerase activity to cancer and aging to date.[49] Human cultured cell samples representing 18 tissue types were assayed for telomerase activity to establish a relationship between telomerase and cell mortality. Ninety-eight percent of immortal cell samples gave positive results for telomerase activity, while none of the mortal cell samples did. The relationship between telomerase and cancer was equally convincing. An

astonishing 90 % of samples made up of 12 different human tumor types showed telomerase activity, while none of the 50 normal somatic tissue samples showed activity.

These results provided a rationale for targeting the telomerase enzyme as a potential treatment against cancer. A strategy of inducing G-quadruplex stabilization in the telomere in order to inhibit telomerase activity and subsequent human tumor growth was envisioned (Figure 1.10).[50] In the presence of G-quadruplex stabilizing ligands, the now folded telomere could be imagined to lose its function as an active substrate for the enzyme. This targeting of the telomerase substrate constitutes an independent mechanism of telomerase inhibition compared to the more classical approach of designing direct inhibitors of the active site. The latter classical approach is exemplified by the design of a complementary strand of peptide nucleic acid that targets the telomerase's RNA template in order to block reverse transcription.[51] This earlier method has now been largely superseded by the pursuit of finding G-quadruplex ligands for sequestering the single stranded 3'-overhang telomere substrate. A proof of principle was achieved when the G-quadruplex interacting agent, BSU1051, was shown to inhibit human telomerase.[52]

Figure 1.10   A potential strategy to develop G-quadruplex ligands into anticancer therapeutics:

The G-quadruplex binding ligand targets the G-rich 3'-overhang of the telomere,

stabilizes G-quadruplex DNA structures, and leads to cell growth arrest in the

tumor cell.
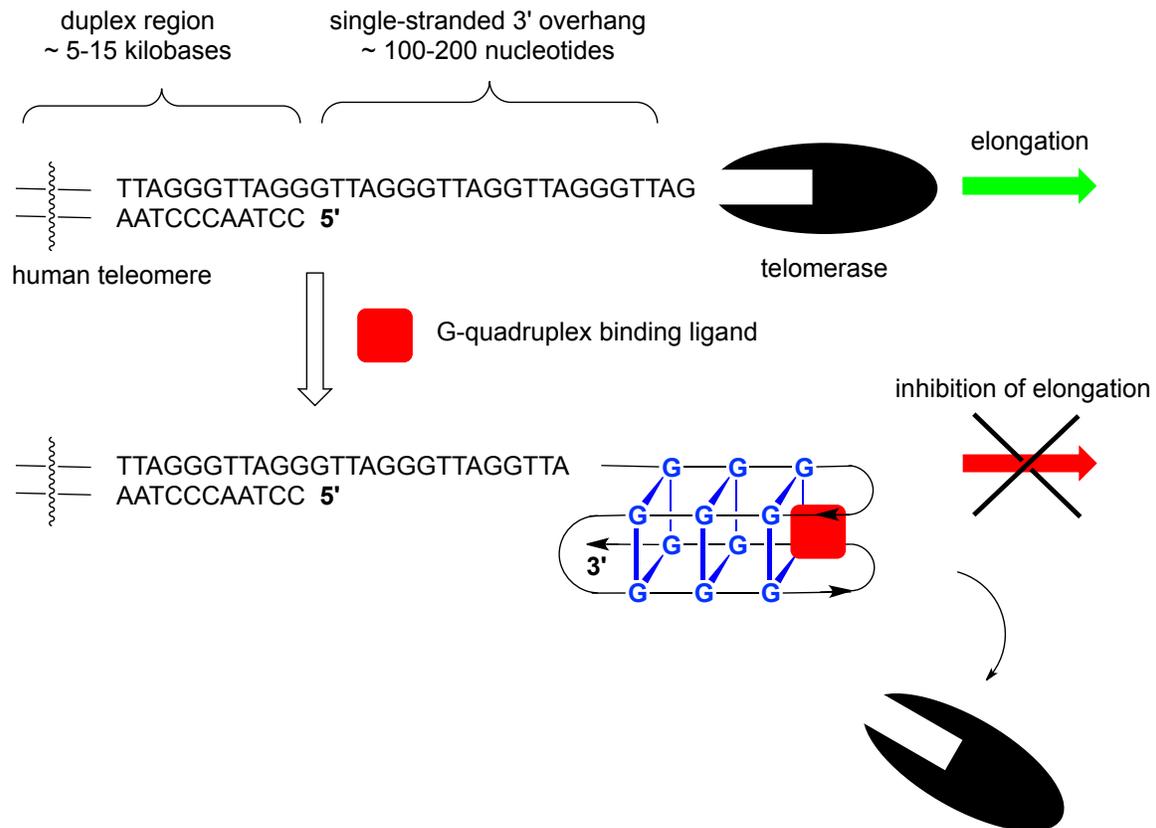
1.5   The Search for G-Quadruplex Binding Ligands

An explosion of interest in developing G-quadruplex ligands as potential

telomerase inhibitors followed the BSU1051 precedent and has kept up for over a

decade.[53-61] Common structural themes are apparent in many of the G-quadruplex ligands (Figure 1.11). A large planar aromatic core provides a favorable surface on which G-

Figure 1.11 Some early G-quadruplex binding ligands derived from anthraquinone, acridine, and perylene structures.
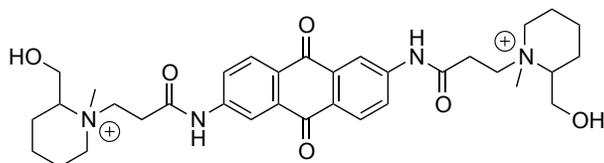
BSU1051
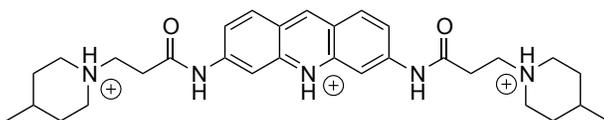
3,6-disubstituted acridine
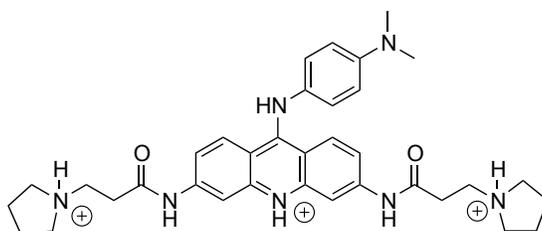
BRACO19
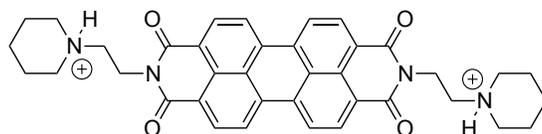
PIPER

tetrads can π-π stack. Flexible side arms that have a potential to hold a positive charge are able to form favorable electrostatic interactions with negatively charged quadruplex phosphate groups. This is often accomplished by the presence of tertiary amines, which will be protonated under physiological pH in aqueous solution, or by the use of quaternary ammonium groups that hold an intrinsic positive charge. Cationic side groups can also enhance water-solubility that may be compromised by the nonpolar hydrophobic core. It is observed that end-stacking of ligands to the exposed face of the terminal G-tetrad is an important mode of interaction for many of the G-quadruplex binding ligands with the quadruplex.

A series of 3,6-disubstituted and 3,6,9-trisubstituted acridine derivatives were developed inside Stephen Neidle's laboratory in addition to the structurally related 2,6-diamidoanthraquinone, BSU1051. Both classes share a tricyclic electron deficient aromatic core with amide linked cationic ammonium groups. A first generation 3,6-disubstituted model compound inhibited telomerase elongation approximately 10 times better than BSU1051, with an IC$_{50}$ of 1.3 μM relative to the IC$_{50}$ of 23 μM found for BSU1051.[62] In the case of BRACO19, another order of magnitude of enhanced telomerase inhibition was afforded after an additional substitution at the 9 position.[63] BRACO19 also demonstrated potent *in vivo* and *in vitro* antitumor activity. A crystal structure of a complex between the acridine, BSU6039, and an *Oxytricha* telomere G-quadruplex showed the ligand to be end-stacked (Figure 1.12).[64]

Another first generation G-quadruplex ligand is the perylene diimide, PIPER.[65] This ligand was shown to quantitatively inhibit telomerase activity at 100 μM concentrations. Spectroscopic evidence also provided information about ligand-

Figure 1.12   Crystal structure of a ligand-G-quadruplex complex depicting a BSU6039 ligand

(magenta) end-stacking to the terminal G-tetrad of an *Oxytricha* telomere G-

quadruplex (PDB entry B1L1) modeled with UCSF Chimera.
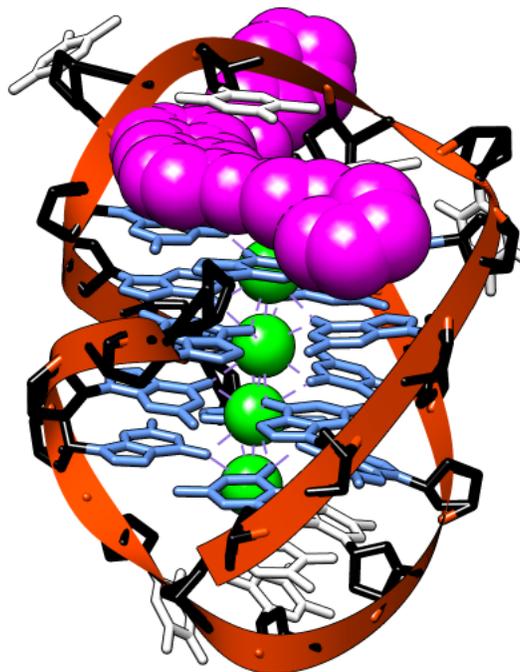
quadruplex complex formation. Both intercalation and terminal guanine quartet end-stacking binding modes were presented for different human telomere derived G-quadruplexes.

Cationic porphyrins are a class of G-quadruplex ligands that have received considerable attention (Figure 1.13). The most heavily studied of these is TMPyP4, which was rationalized as a potential G-quadruplex ligand due to it having a planar structure with comparable size to the G-tetrad as well as having cationic quaternary aromatic

Figure 1.13   Some G-quadruplex binding ligands based on the porphyrin structure.

TMPyP4

TMPyP3

3,4-TMPyPz
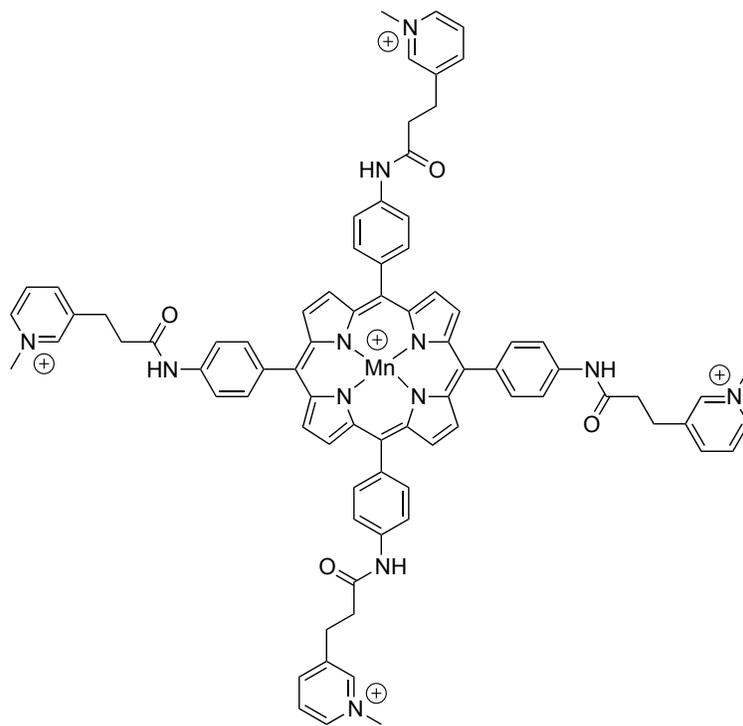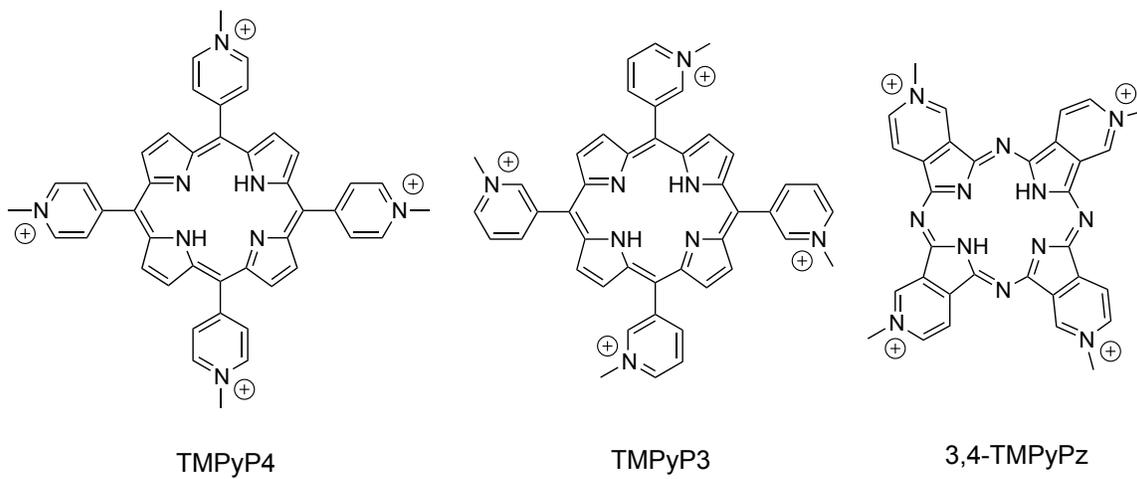
pentacationic manganese (III) porphyrin

ammoniums.[66,67] Binding constants as high as 10^6 M^{-1} have been determined from optical studies,[68] and a telomerase IC_{50} of 6.5 μM was measured.[67] These ligands initially showed good selectivity over duplex DNA, and this selectivity has been dramatically improved in later generations of porphyrins.[66] For example, the pentacationic manganese (III) centered metalloporphyrin is a 10000 fold increase more selective for G-quadruplex than duplex DNA.[69] Preference for binding to different G-quadruplex strand topologies is also of interest in the context of the development of cancer therapeutics since different topologies may be involved *in vivo*. The analogue, TMPyP3, was shown to exclusively promote G-quadruplex formation in parallel stranded structures over the TMPyp4 and TMPyP2 variants.[70] In contrast, the porphyrazine, 3,4-TMPyPz can induce the formation of antiparallel human telomere G-quadruplexes.[71] Debate existed in the early literature over whether the preferred mode of interaction for TMPyP4 was intercalative or end-stacking but this debate has largely been resolved with the conclusion that both modes are important. Phan reported the NMR based solution structure of a complex with a c-myc promoter G-quadruplex that showed an exclusive end-stacking arrangement (Figure 1.14).[72] Alternatively, Wei *et al.* reported that end-stacking was favored for a parallel stranded G-quadruplex topology, but that intercalation was favored for dimer hairpin and intramolecularly folded topologies.[73]

Known ligands that bind to G-quadruplex DNA are not limited to synthetic molecules, and in fact, one of the most potent telomerase inhibitors, telomestatin, belongs to the natural product family (Figure 1.15). Telomestatin isolated from *Streptomyces anulatus* is an uncharged macrocycle of seven consecutive oxazole rings and a lone thiazoline ring.[74] It has been shown to have one of the lowest IC_{50} against telomerase at

Figure 1.14   Solution structure of a ligand-G-quadruplex complex depicting a TMPyP4 ligand

(magenta) end-stacking to the terminal G-tetrad of a human c-myc promotor

G-quadruplex (PDB entry 2A5R) modeled with UCSF Chimera.
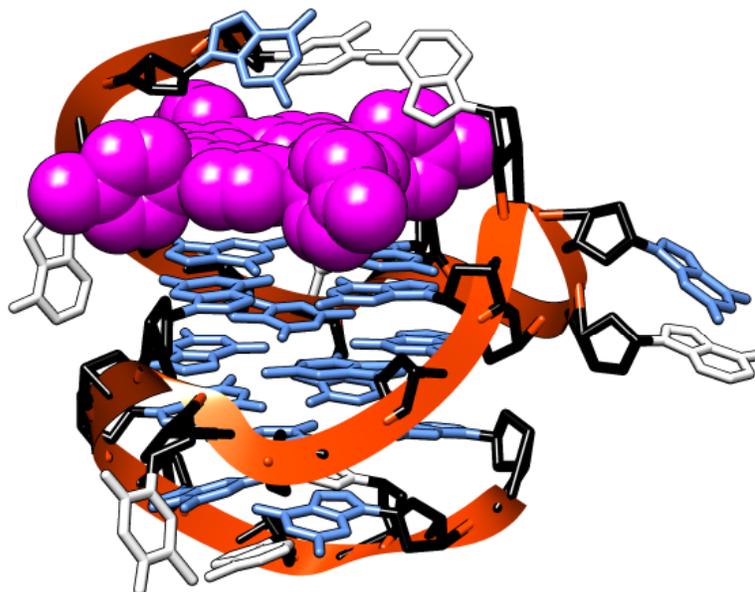
0.005 μM. Furthermore, no measurable inhibition of DNA polymerase or reverse transcriptase was found. Proof of interaction with an antiparallel intramolecular human telomeric G-quadruplex[75] and of inhibition of cell growth towards human leukemia cells[76] has been demonstrated. Of note, is that a multi-step convergent total synthesis of telomestatin has been reported and was based on the strategy of forming the hydrolytically labile thiazoline ring during the final stage.[77] Total synthesis of a natural product like telomestatin represents synthetically a significantly more demanding task than that for conventional synthetic ligands.

Figure 1.15 Some natural product and natural product analogue derived G-quadruplex

binding ligands.

telomestatin   berberine   coralyne
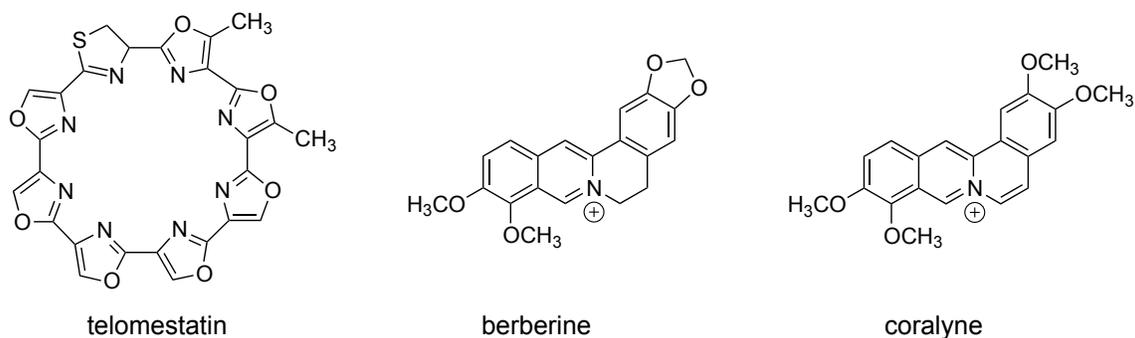
The plant alkaloid berberine is an additional natural product of interest for the development of anticancer agents. The singly charged five-fused ring molecule is a known moderate inhibitor of telomerase.[78] Derivatization and aromatization to the semi-synthetic analogue, coralyne, resulted in improvements in telomerase inhibition, G-quadruplex affinity, and G-quadruplex selectivity over double-stranded DNA with respect to berberine.[79]

All but a few of the IC_{50} values towards human telomerase reported for the various ligands were measured by the telomeric repeat amplification protocol (TRAP) assay first developed by Kim and coworkers[49] or by a modified TRAP assay such as Gomez's.[80] De Cian and colleagues have proposed a reevaluation of the inhibition of telomerase due to intrinsic flaws of the TRAP assay.[81] The telomerase inhibitory action of G-quadruplex ligands may be exaggerated in the polymerase chain reaction (PCR) based TRAP assay compared to direct measure methods, because G-quadruplex ligands are more capable inhibitors of telomere amplification rather than elongation. This

suggests that the mechanism of action of G-quadruplex ligands to arrest cell proliferation in many immortalized and tumor cells may involve a fundamentally different mechanism less dependent on the telomerase enzyme.

1.6   Towards Template-Assembled Guanine Tetrads and Guanine Quadruplexes

The concept of using a covalent template to preorganize multiple nucleic acid elements for intramolecular and intermolecular self-assembly has its origins in the simple linkage of nucleobases to cyclodextrins, crown ethers, and polycyclic aromatics (Figure 1.16). Generation of an intramolecular hydrogen bonded complementary base pair from a β-cyclodextrin bifunctionalized with thymine and adenine nucleobases was reported by Nagai.[82] Adenine's propensity to form intramolecular base pairs with thymine and uracil has also been investigated by the construction of "DNA boxes" from alkyl linked crown ethers and by the synthesis of artificial dinucleotide complexes from palladium cross couplings of protected nucleosides with polycyclic templates.[83-86]

The driving force for template-assembled synthesis of nucleic acid elements, such as for the previous examples of synthetic hydrogen bonded base pairs, lies in the thermodynamic principle of entropy, and is analogous to the concept used in the creation of synthetic templated proteins.[87] Covalent templates function to remove translational and rotational degrees of freedom from the free recognition elements thereby making the

Figure 1.16   Template-assembled thymine adenine base pairs: a) intramolecular with

cyclodextrin template, b) intermolecular with crown ether template forming "DNA

box", and c) intermolecular with artificial dinucleotide complex.
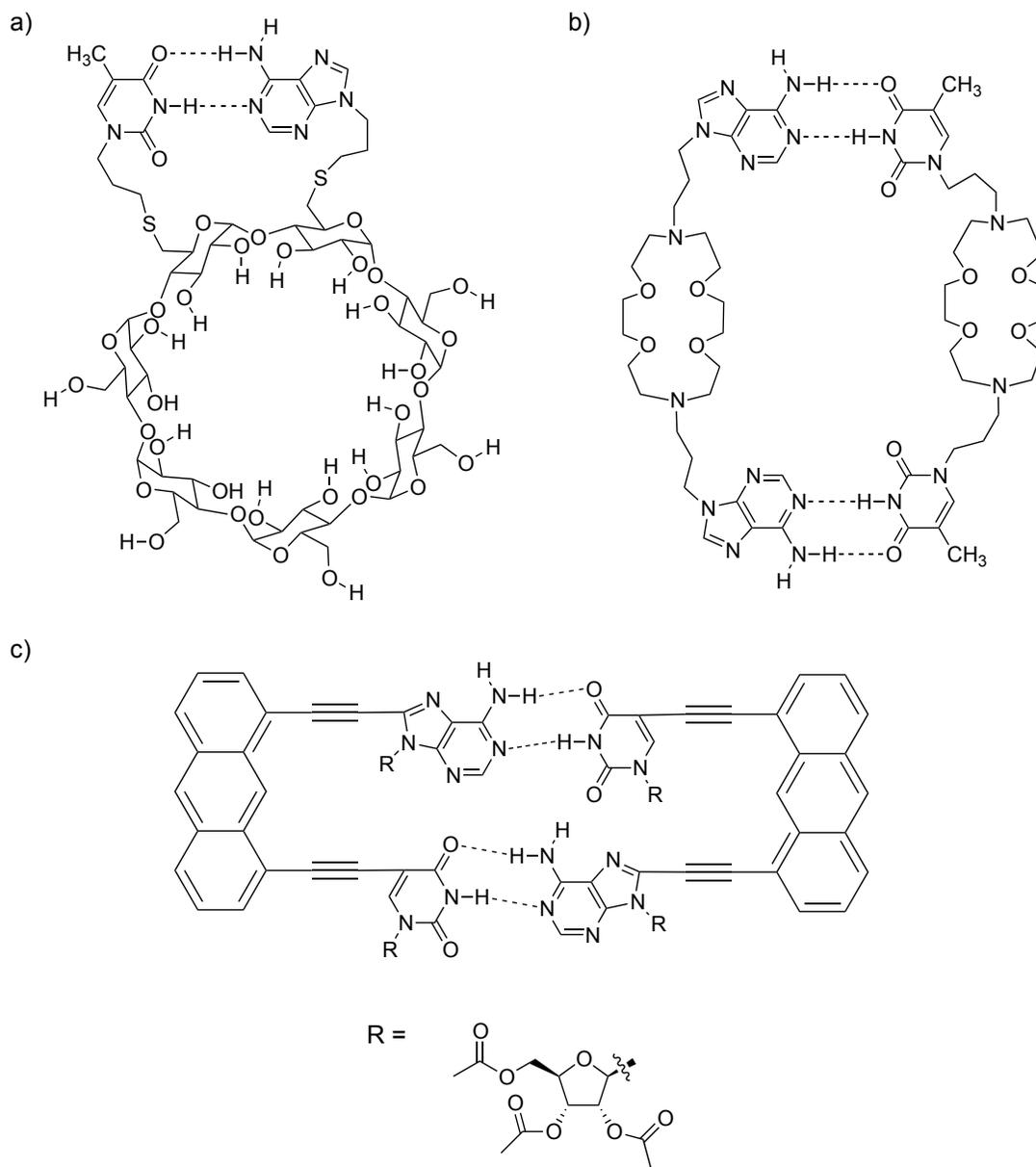
absolute value of the entropically disfavored assembly process smaller for a templated system as compared to a template-free system. A thermodynamically stabilized templated system results, whereas self-assembly of the free recognition elements would not be expected to be thermodynamically feasible.

Calix[4]arene templates have received considerable attention for the self-assembly of nucleobases. Deoxythymidine residues have been attached via phosphate linkages (Figure 1.17).[88] The resulting bis-coupled nucleotide-calixarene conjugate was proposed to self-assemble through thymine-thymine mediated hydrogen bonding into a termolecular supramolecule. Davis's laboratory has investigated the functionalization of calix[4]arene-1,3-*alternate* scaffolds with modified guanosine nucleosides in an effort to design artificial ion channels (Figure 1.18).[89,90] Cation stabilization of intermolecular G-quartets facilitated the creation of nanotubes and a lipophilic ion-pair receptor from these constructs.

Figure 1.17   a) A di-functionalized deoxythymidine calixarene containing phosphate linkages.

b) Proposed thymine-thymine hydrogen bonding of trimer assembly.
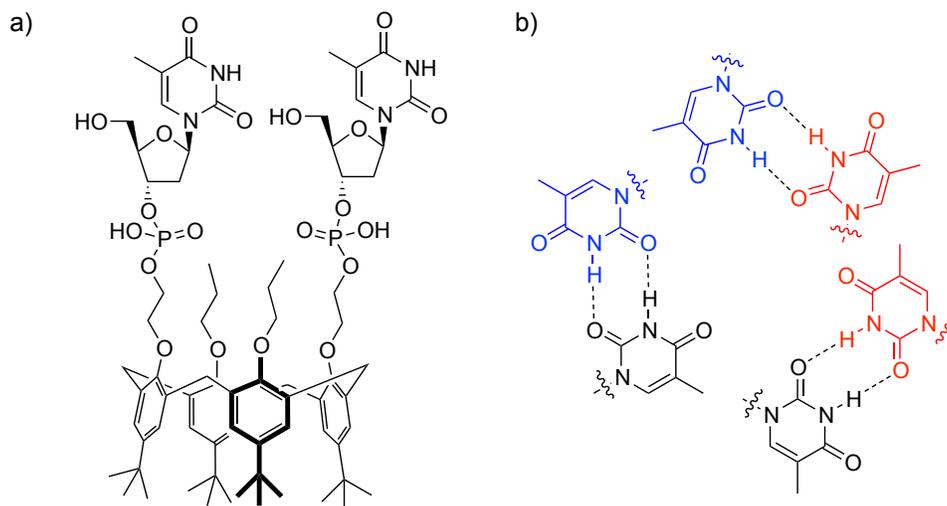
Figure 1.18   A guanosine functionalized calixarene with alternate scaffold assembles into

intermolecular G-quartet based nanotubes in the presence of sodium.
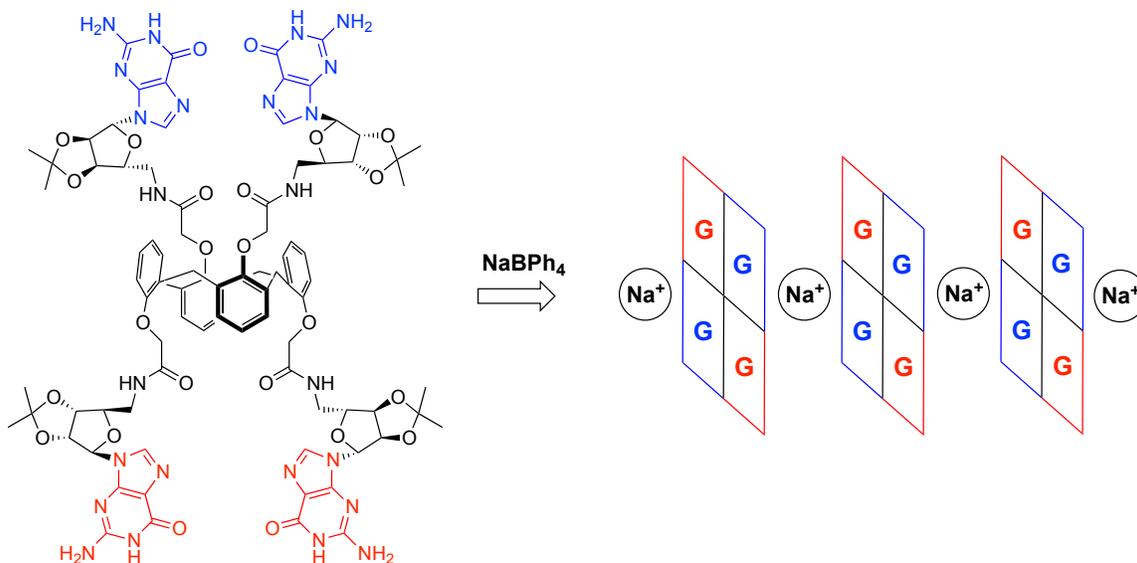
The cavitand, another prevalent structure used by synthetic chemists to control non-covalent interactions, has found use in the synthesis of a template-assembled cation-free G-quartet (Figure 1.19).[91] In comparison to Davis's calixarene template, the rigid cavitand facilitated preorganization on one face of the template. Intramolecular G-quartet formation was observed accordingly. In the presence of cesium cation, tuning of the unimolecular species into dimeric structures was afforded.[92] It should be noted that the guanosine coupled calixarenes and cavitands described thus far share a common chemical limitation in that they are lypophilic chloroform soluble systems. It is a goal of the synthetic chemist interested in the bottom-up design of minimal models of telomere-like G-quartets and G-quadruplexes to impart water-solubility on such systems. Water-soluble

Figure 1.19   A cavitand based cation-free template-assembled G-quartet in chloroform that

dimerizes in the presence of cesium cation.
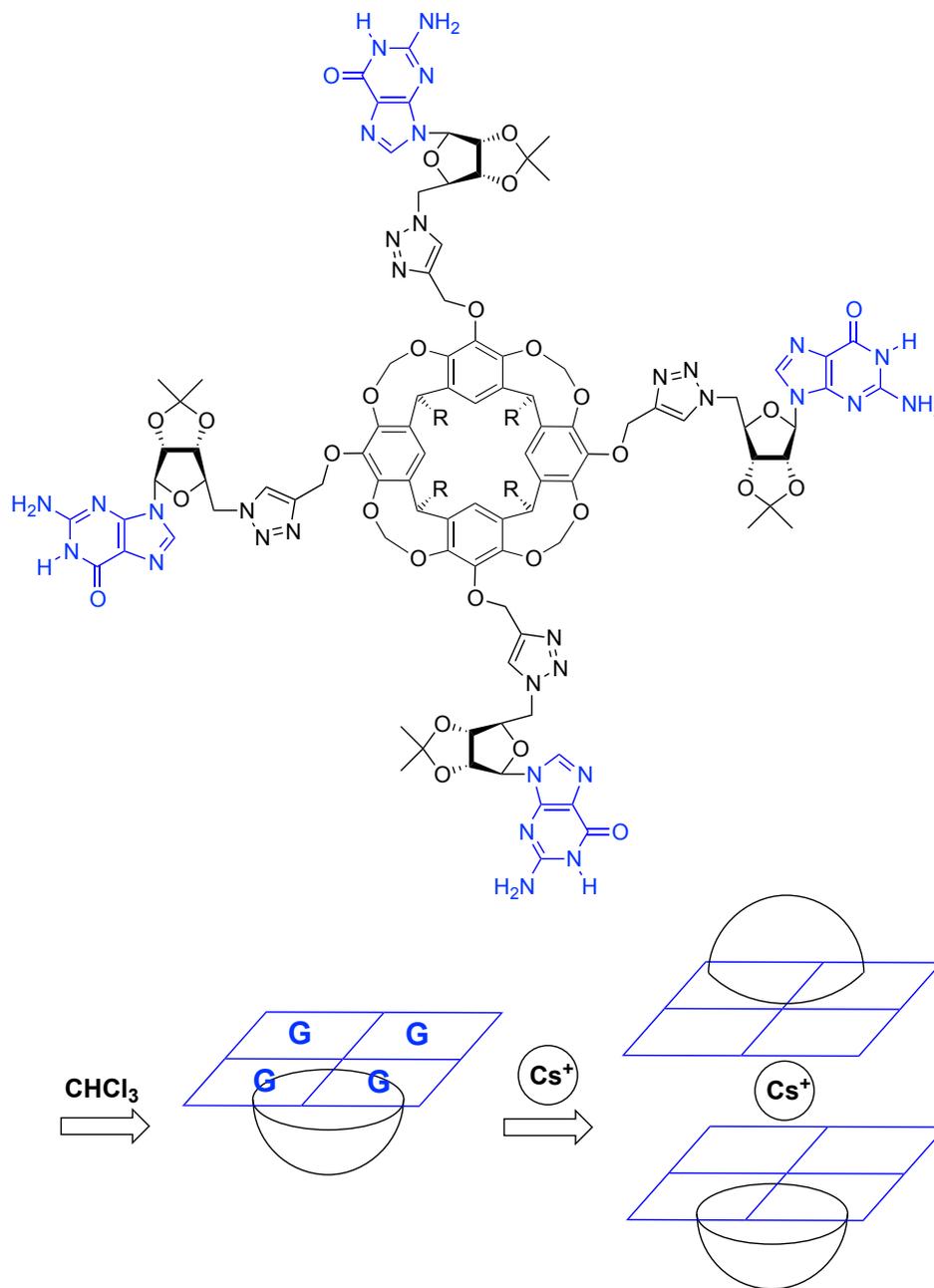
synthetic guanine tetrads may be beneficial to the understanding of *in vivo* processes involving the telomere and may aid in the discovery of the anticancer potential of G-quadruplex binding ligands.

The role of templates in G-quadruplex formation was first explored with the solid-phase phosphoramidite synthesis of tetra-end-linked oligonucleotides (Figure 1.20).[93,94] Control of G-quadruplex topology with the use of non-nucleotidyl linkers of varied length and orientation proved to be ineffective as all template G-quadruplexes preferred a parallel strand association.[95] Tetra-end-linked G-quadruplexes demonstrated application as an anti-HIV aptamer.[96] A parallel template assembled G-quadruplex was also achieved after the oxime-based coupling of G-rich oligonucleotides to Mutter's cyclic peptidic scaffold.[97] Various G-quadruplex ligands were shown to interact with this peptidically templated G-quadruplex.[98]

Figure 1.20   Examples of template-assembled G-quadruplexes based on a) a tetra-end-linked template and b) Mutter's template.

a)

non-nucleotidyl tetra-end linker

b)
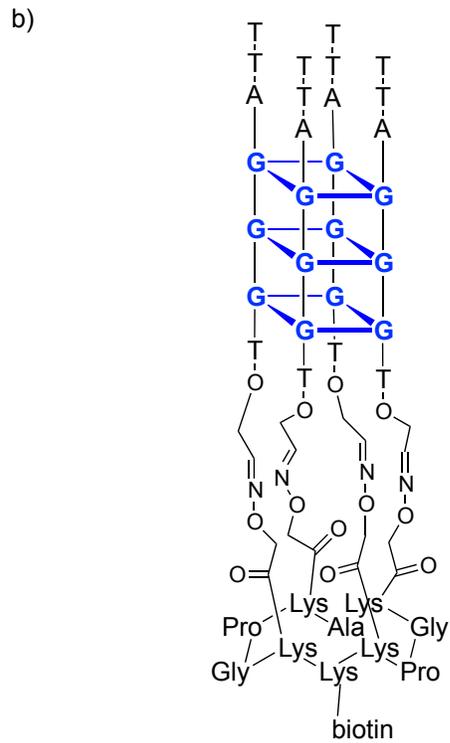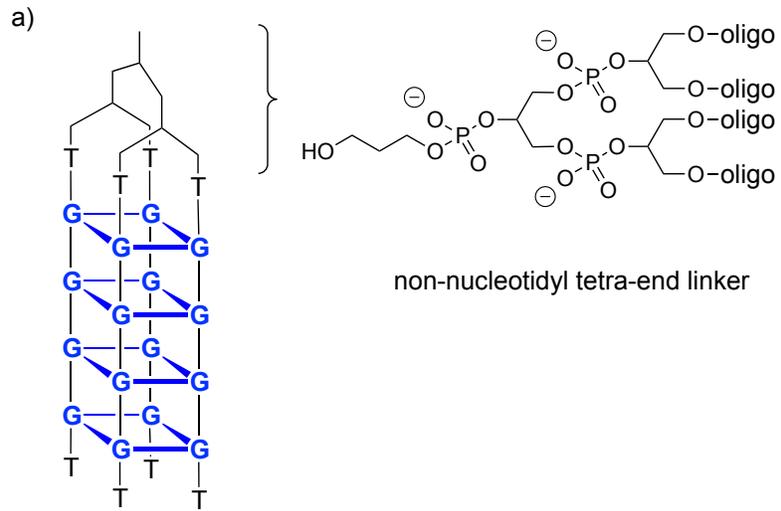
1.7   Canonical Nucleobase Quartets not Based on Guanine

In relation to guanine, hydrogen bonded tetrads or quartets built from the other canonical nucleobases of DNA or RNA are rare. Of those reported, uracil tetrads are the most frequently encountered (Figure 1.21). The first being the report of a U-quartet terminating the parallel tetramolecular RNA G-tetraplex of r(UGGGGU) strand sequence.[99] The existence of such a novel structure was hinted at from the observation of a downfield uracil imino signal in the ^{1}H NMR spectrum in water. Crystal structures have corroborated the assembly of U-quartets in the same RNA tetraplex[100] and from the free base when coordinated in a platinum complex.[101] Tetrads of uracil in all but a few cases belong to nucleoside residues in G-quadruplex forming oligonucleotide strands. Thus, it may be assumed that G-quadruplex scaffolds must be U-quartet stabilizing, however the opposite case has been shown to be true.[102] A cavitand template-assembled synthetic U-quartet has also been demonstrated.[103] Similar to the uracil tetrad imbedded quadruplexes, is a G-quadruplex structure terminating in a thymine quartet, which had been deduced by NMR studies.[104] Several reports exist proposing stabilized A-tetrads as part of adjacent G-quadruplex structures.[105,106] Evidence for a bona fide cytosine hydrogen bonded tetrad is lacking.

Figure 1.21   A hydrogen bonded planar cation bound uracil tetrad and cation free thymine

tetrad

U-tetrad   T-tetrad
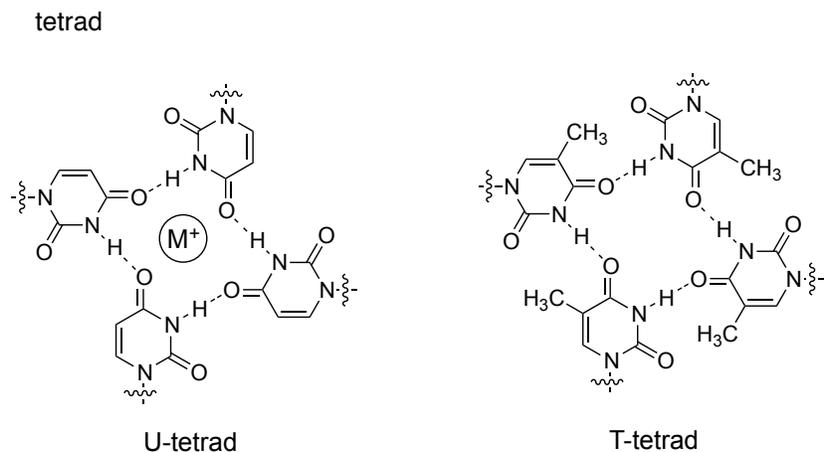
1.8   Thesis Aims

The global aim of this thesis work was established almost five years ago with the concept of connecting four guanine nucleosides onto a cavitand substrate in order to create a molecule that was not only water-soluble, but that could template assemble a guanine quartet in an aqueous environment (Figure 1.22)! At the time that this project was undertaken, there were no known precedents of synthetic methodologies used for four-fold covalent couplings of guanine nucleotides on a single molecule. Nor were there even any known water-soluble substrates bearing four guanine bases. It was an intention from the start to make the potential water-soluble template-assembled guanine-cavitand construct to be the most DNA-like possible. This intention served two purposes. First, it narrowed down the synthetic strategy used for the target molecule to those tried and true reactions pioneered for the chemical synthesis of DNA, like the phosphite triester and

phosphoramidite reactions used herein, and secondly, it sought to improve the *in vivo* significance of the final construct. In the context of a bottom-up synthesis of a minimal model of telomere DNA, working in Nature's solvent and making use of proper DNA elements like the correct phosphate linkages seemed advantageous.

Questions arose as work progressed, and these questions were used as a guide to further research. Is it possible to solubilize four guanosine residues and a template in water? What water-solublizing groups should be chosen, how many groups are needed, and where will they go? Are there reactions and templates chemically efficient enough to couple four residues with? How will the complex mixture of diastereomers formed after the coupling synthetic step be handled, deblocked, characterized, and finally purified to arrive at the pure target molecule? Could a G-quartet exist in water in the absence of cation stabilization? Could a free unimolecular G-quartet face be created, and could this face act as an artificial receptor for G-quadruplex binding ligands? How energetically feasible is it to form template assembled nucleobase quartets of the other canonical bases? Could these systems described here have practical application towards the study of G-quadruplex binding ligands as novel anticancer therapeutics? It is the hope of the author that solutions to questions like these will have been provided in the following chapters.

Figure 1.22   Aims of this thesis work included: 1. developing DNA synthetic methodologies for
the construction of a water-soluble template-assembled G-quartet, 2. probing
intermolecular association, cation binding, and G-quadruplex ligand binding, and
3. exploring analogous non-guanine nucleobase systems.

cavitand template

deoxyguanosine

water-solublizing groups?

cation bound G-quartet?

DNA synthesis???

cation free G-quartet?

other nucleobase quartets?

G-quadruplex ligand receptor

G-quartet dimerization?
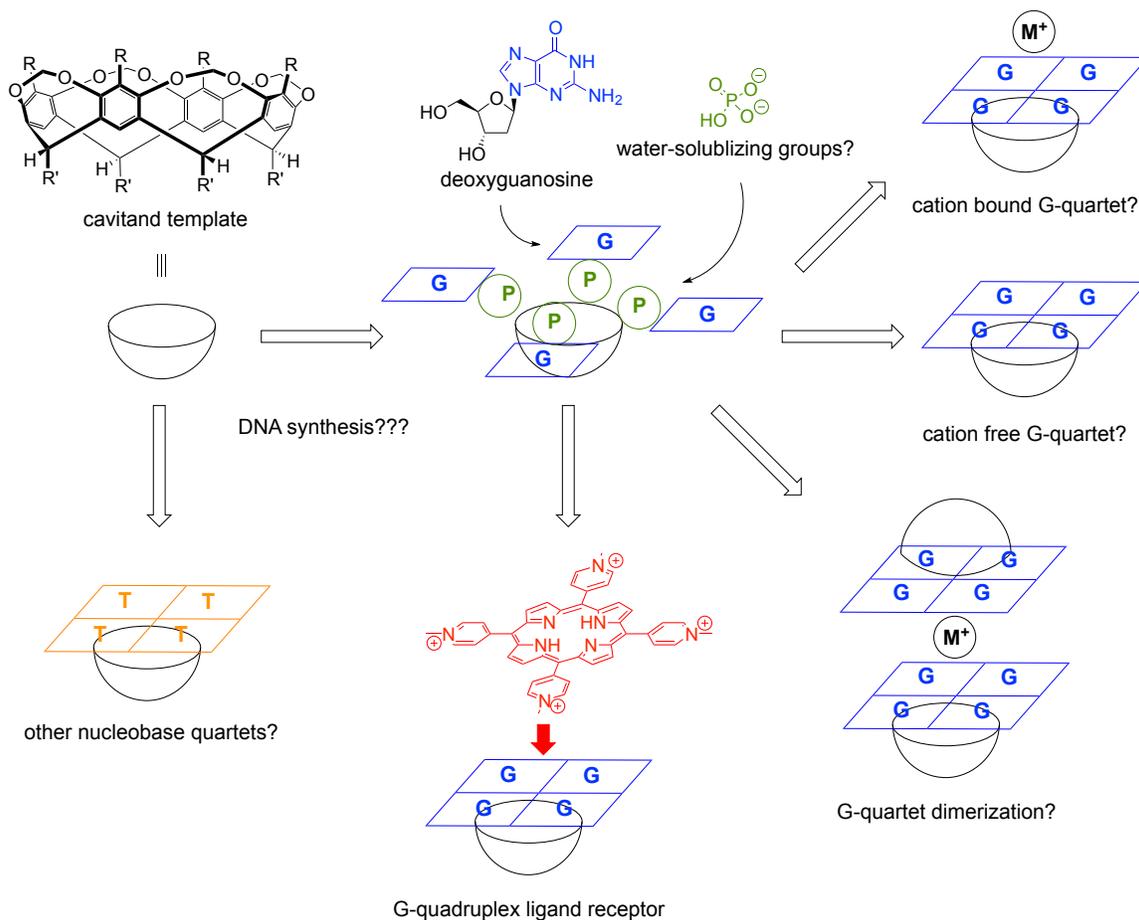
Chapter 2: Synthesis of Phosphate-Linked Cavitand-Nucleotide Conjugates

2.1   Synopsis

The phosphite triester methodology for DNA synthesis was applied in the construction of tetra-coupled guanine based cavitand-nucleotide conjugates linked through natural 3' phosphate diesters. This synthesis represents an early example of a water-soluble substrate covalently bound with four guanine bases and the first template coupled to four deoxyguanosine residues through natural phosphate linkages. A novel alcohol rim functionalized cavitand was used in the construction of one of the cavitand-nucleotide conjugates. The phosphoramidite methodology for DNA synthesis was similarly applied. A class of silyl ether protected phosphoramidite reagents of the canonical deoxynucleoside series was developed for use in the synthesis of tetra-coupled cavitand-nucleotide conjugates. In addition, the phosphoramidite approach allowed a route to the synthesis of base-protected cavitand-nucleotide conjugates. The synthesis herein was designed with the goals in mind of synthesizing a water-soluble template-assembled guanine quartet as a potential simplified model of G-quadruplex telomere DNA or as an artificial G-quadruplex binding ligand receptor and of discovering non-guanine based template assembled nucleobase quartets.

2.2 Synthetic Strategy

A strategy was envisioned towards the synthesis of a water-soluble template assembled guanine quartet. Various molecular structural requirements were incorporated into this strategy with the foresight that such a template assembled system could act as a minimal design for telomeric G-quadruplex DNA or as a receptor for G-quadruplex binding ligands. The [4]cavitand was selected as a suitable supramolecular template (Figure 2.1). The bowl shaped cavitand is rigidified through acetal bridges preventing

Figure 2.1   Various depictions of alcohol functionalized cavitands.
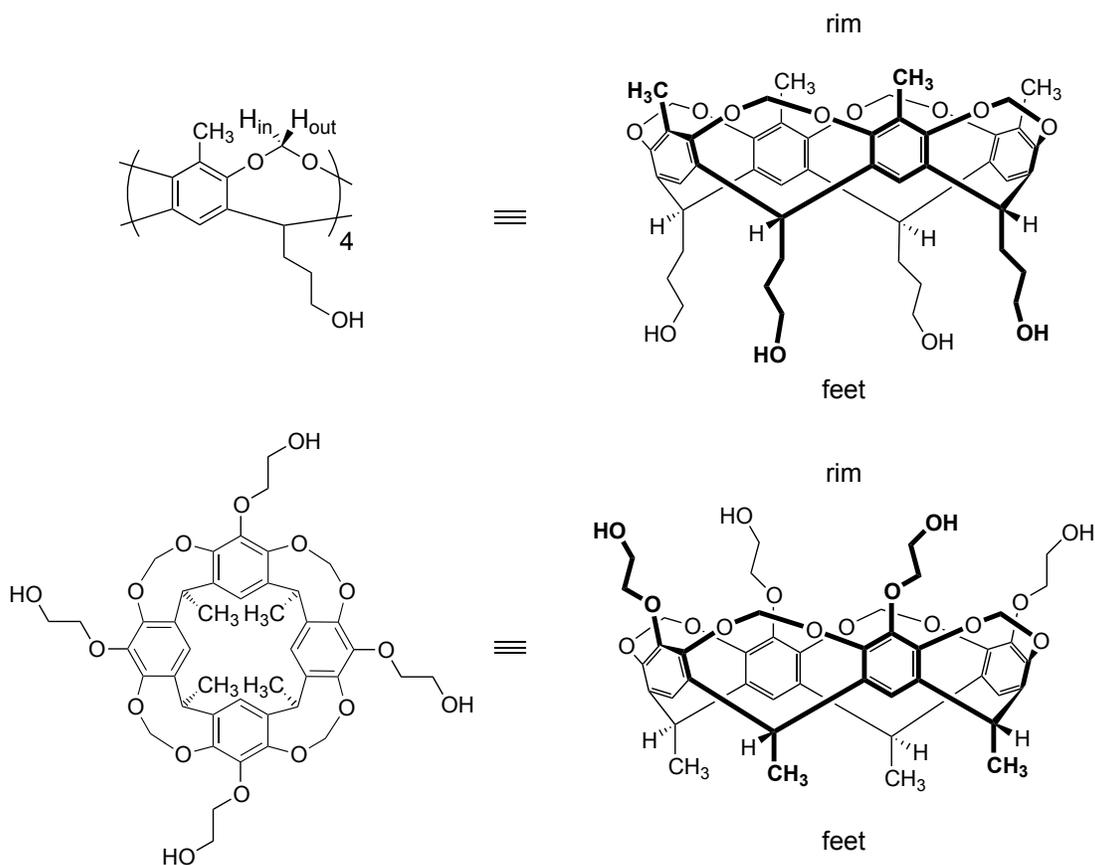
undesirable bond rotations and flexibility and possesses a four-fold proper axis of rotation symmetry element like that found in the planar G-tetrad. The ability of a cavitand to pre-organize a guanine and an uracil nucleoside quartet on one face has been demonstrated in a nonpolar organic solvent.[103,91]

Coupling four guanine nucleotides through phosphate linkages to a single substrate had been unprecedented, and already established methodologies for the chemical synthesis of DNA were drawn upon as a means to accomplish this synthetic challenge. Application of existing methodologies would be advantageous in creating a target molecule that would have high "DNA-like" character by incorporating 2'-deoxynucleosides attached through a potentially water-solubilizing natural 3' phosphate diester linkage. Solubilizing four guanine residues on an individual template in an aqueous medium had also not been demonstrated when work began.

Landmarks in the chemical synthesis of DNA include Michelson and Todd's first synthesis of an internucleotide bond, Khorana's phosphate diester based total synthesis of the structural gene for a yeast alanine transfer RNA, Letsinger's development of the phosphite triester methodology, and Beaucage's preparation of phosphoramidite reagents and their use in modern automated solid-phase DNA synthesizers (Figure 2.2).[107-111] All but Khorana's method for 3',5' phosphodiester bond construction utilize the more chemically reactive mode of coupling whereby a displacement reaction takes place with 5' primary alcohol nucleophile at an activated 3' secondary phosphorus III or V center. Thus chemical methods for internucleotide bond formation tend to be in opposition to *in vivo* enzymatic bond formation, which involve condensation of a 3' secondary alcohol nucleophile with a 5' triphosphate nucleoside.

41

Figure 2.2   a) Michelson and Todd's synthesis of a thymine thymine dinucleotide. b)
Khorana's phosphate diester coupling methodolgy utilized in first total gene
synthesis. c) Letsinger's phosphite triester coupling approach. d) Beaucage's
original phosphoramidite procedure. Protecting groups: Ac: acetyl; An: anisoyl;
Bn: benzyl; Bz: benzoyl; DMT: 4,4'-dimethoxytrityl; MMT: 4-monomethoxytrityl;
TCE: 2,2,2-trichloroethyl.
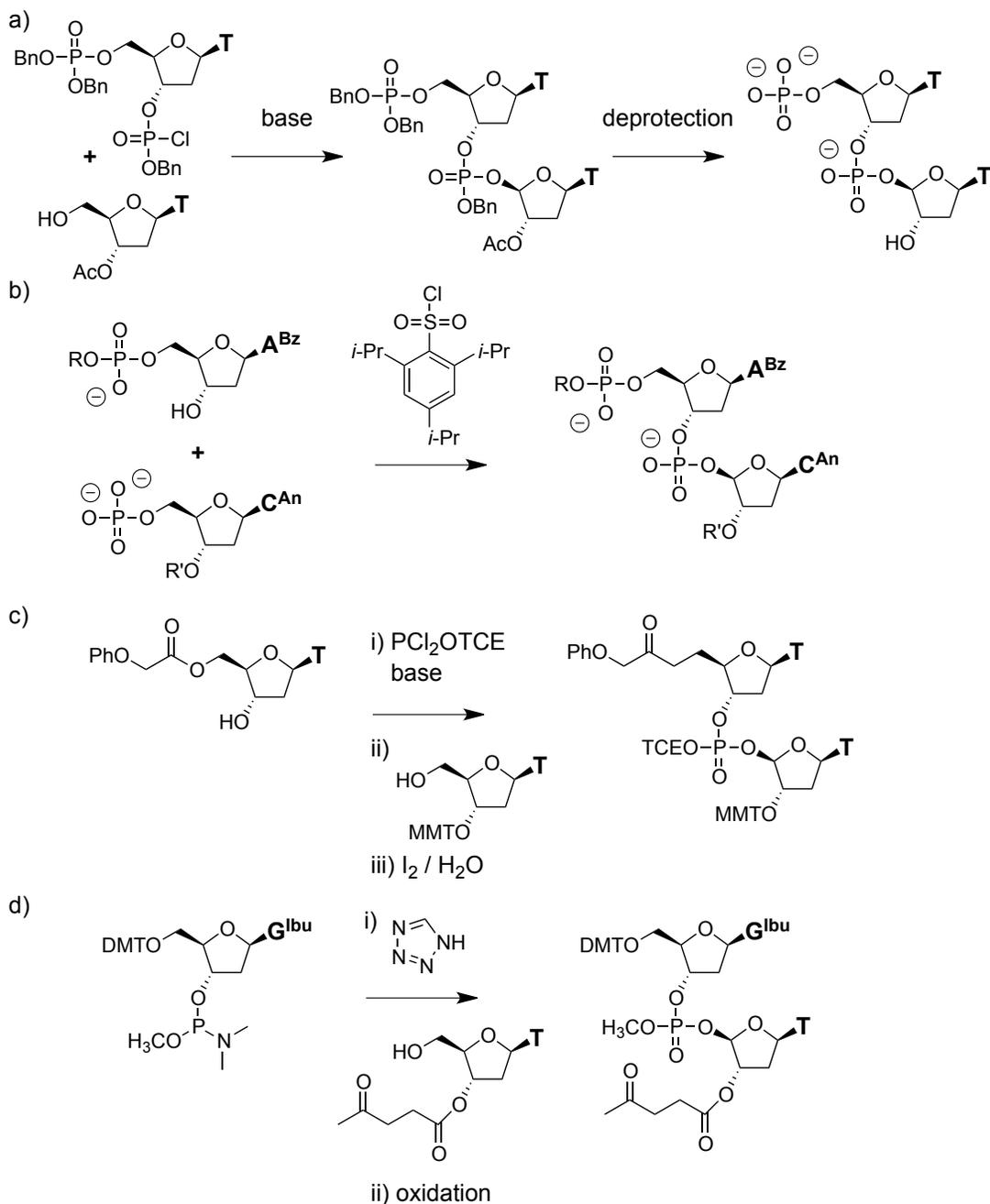
2.3 Phosphite Triester Synthesis of Guanine Base Phosphate-Linked Cavitand-Nucleotide Conjugates

Significantly improved yields and shortened reaction times were reported by Letsinger for the use of phosphorus III containing phosphorochloridites as coupling agents in the formation of 3',5' phosphate triesters in comparison to the analogous phosphorus V containing phosphorochloridates (Figure 2.2c).[109] In this manual solution-phase phosphite triester approach, an *in-situ* displacement of two chlorides of a phosphorodichloridite by successive addition of a 3' and then a 5' unprotected nucleoside takes place cleanly at -78 ºC in tetrahydrofuran in the presence of a base on the minute time-scale. Brief iodine catalyzed oxidation of the phosphite triester gives the protected phosphate triester analogue. Subjecting the fully coupled and protected oligodeoxynucleotide to deblocking conditions and purification facilitates the completion of the synthesis. The use of simple starting materials, the avoidance of any preparation and isolation of unstable activated intermediates, and the ability to perform two couplings in one pot contribute to the attractiveness of this approach.

Known 2'-deoxyguanosine nucleoside, **1a**, protected at the base amino position as the isobutyryl (Ibu) amide and at the 5' alcohol position with *tert*-butyldimethylsilyl (TBS) ether was chosen to couple together with known cavitand **2** (Scheme 2.1).[112-115] Cavitand **2** possesses primary propyl alcohol feet capable of mimicking the 5' primary alcohol nucleophile required for DNA synthesis. Conventional DNA synthesis utilizes a bulky

Scheme 2.1 Phosphite triester synthesis of guanine base cavitand-nucleotide conjugate at the feet position.
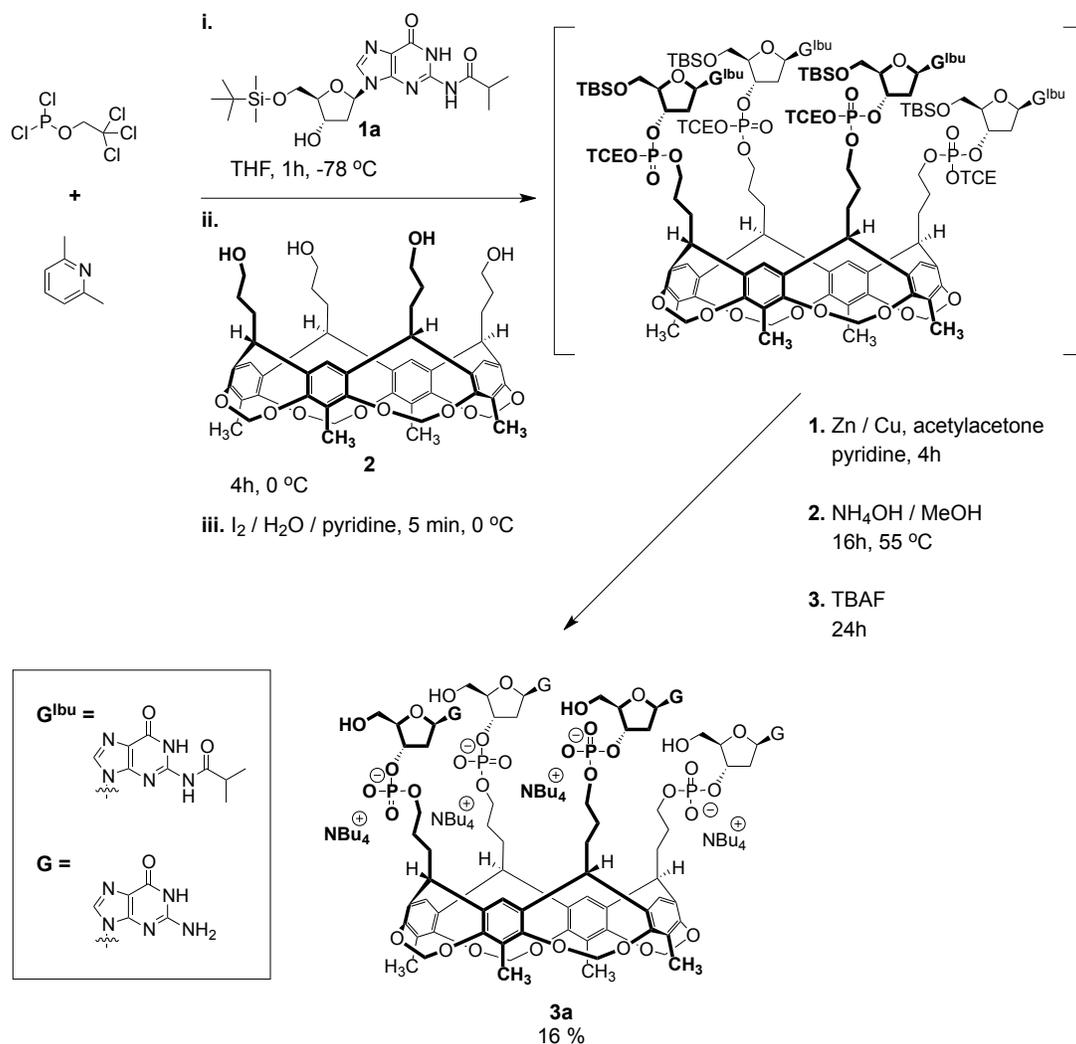
acid labile dimethoxytrityl (DMT) protecting group on the 5' alcohol.[116] For the reasons of minimizing steric demand during four fold coupling and avoiding acid catalyzed depurination side reactions, the TBS group was chosen in favour of the DMT group for 5' protection. In order to achieve complete conversion of starting material **2** to the fully

coupled and protected intermediate conjugate, four equivalents of **1a** per alcohol function was used. Starting material **2** or any mono-, di-, or tri-coupling intermediates could not be detected by thin-layer chromatography after several hours of reaction progress. Coupling and deblocking reactions were carried out in succession without isolation or characterization of the intermediates due to the problematic generation of diastereomeric mixtures when asymmetric phosphate triesters are formed and due to issues of solubility. Deblocking of the 2,2,2-trichloroethyl (TCE) phosphate protecting group was accomplished by zinc / copper couple in pyridine in the presence of acetylacetone.[117,118] Guanine base isobutyryl group deprotection was carried out with ammonia treatment in a water methanol mixture. Final removal of the TBS ether group with TBAF in tetrahydrofuran followed by purification on C_{18} reversed-phase silica afforded the cavitand-nucleotide conjugate **3a** as a tetrabutylammonium salt in 16 % yield from the coupling step. The incorporation of a *n*-tetrabutylammonium counterion into the final purified product was not by design. However, choosing standard silyl ether protection and fluoride cleavage was fortuitous, since this counterion imparted the desired solubility properties on the target while other counterions could not be found to do so. Compound **3a** was observed to have millimolar solubility in the polar protic solvents methanol and water.

Cavitands have a nonpolar concave surface or hydrophobic cavity that may be expected to influence the solubility behavior of a particular cavitand-based construct in water. Thus, an additional goal existed to synthesize a cavitand functionalized at the rim with primary alkyl alcohols. This was accomplished by substitution of known tetrol cavitand **4** with ethyl bromoacetate in the presence of potassium carbonate base followed

by reduction of the ester with lithium aluminum hydride to yield the new tetra-alcohol cavitand **6** (Scheme 2.2).[119,120] Accordingly, cavitand **6** was utilized as a substrate for the phophite triester synthesis of cavitand-nucleotide conjugate **7** and demonstrated identical solubility character to its feet functionalized analogue **3a** (Scheme 2.3).

Scheme 2.2 Synthesis of alcohol functionalized cavitand at rim position.
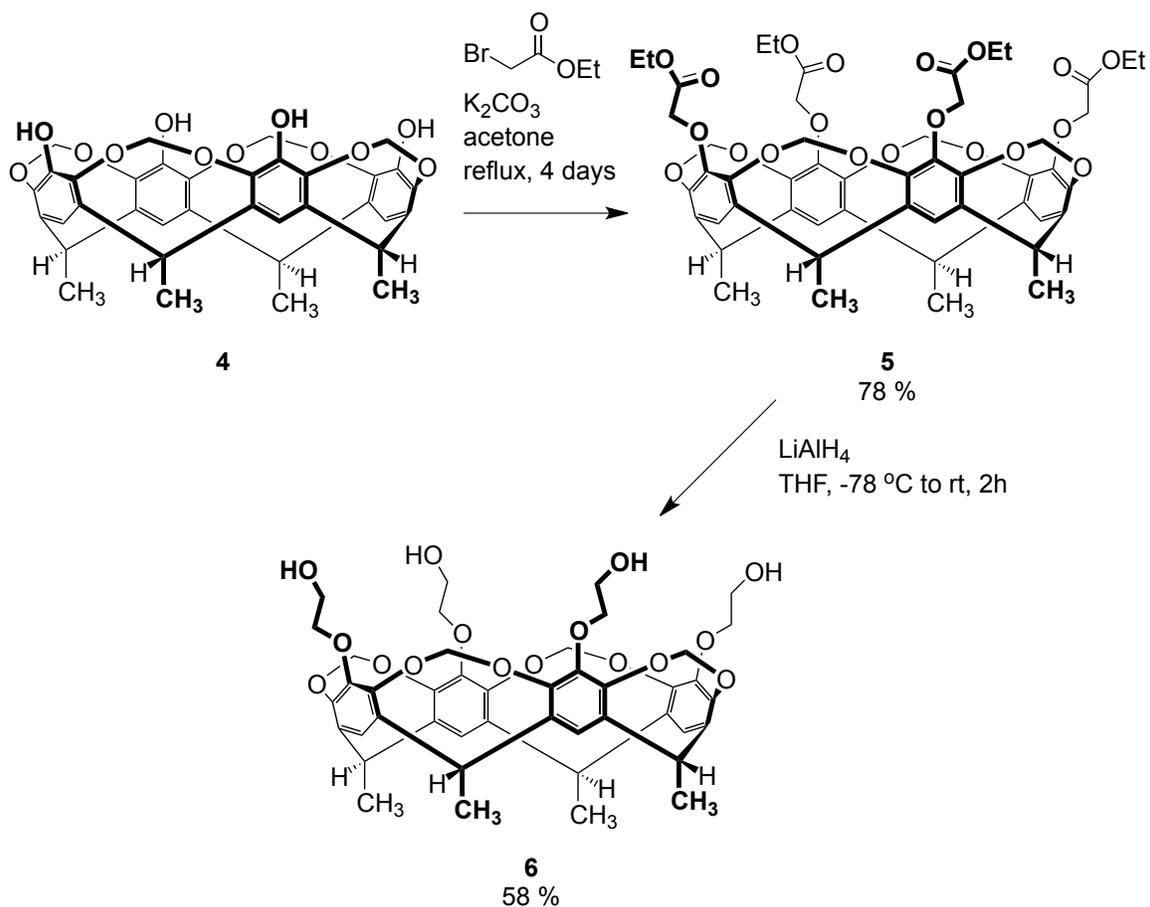
Scheme 2.3 Phosphite triester synthesis of guanine base cavitand-nucleotide conjugate at the rim position.

7
26 %
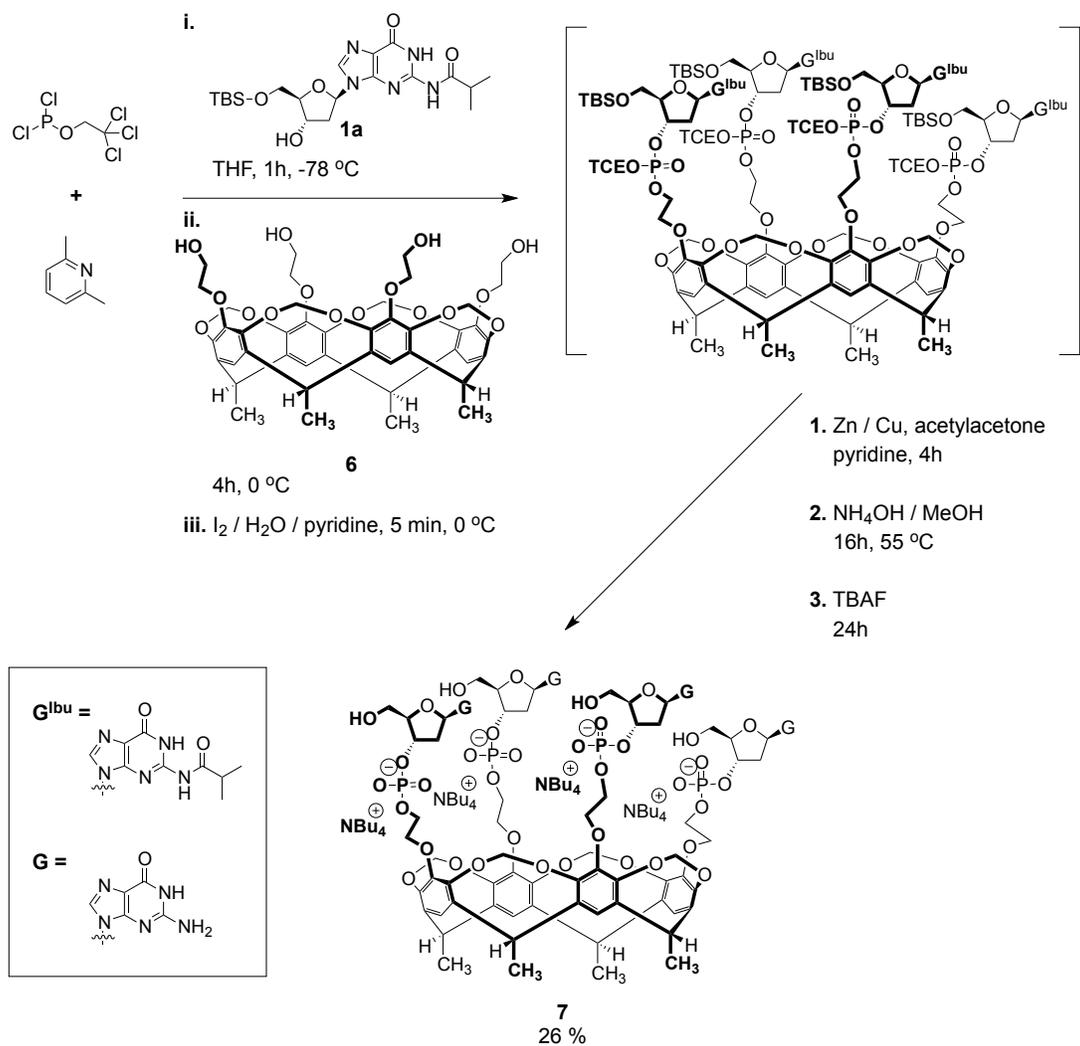
2.4 Phosphoramidite Synthesis of Phosphate-Linked Cavitand-Nucleotide Conjugates of the Canonical Nucleobase Series

Application of phosphoramidite reagents to the synthesis of DNA is widespread and is routinely observed in modern automated solid-phase DNA synthesis or "gene machines".[111] An impressive transformation in the biological sciences has taken place as a consequence of the advent of gene machine technology. It is now possible for a technician who is not a specialized organic chemist to produce a purified oligodeoxynucleotide of approximately 20 residues long in a matter of hours. In the past, such a synthetic feat like that seen in the original construction of a non-repeating icosanucleotide in Khorana's laboratory by manual methods would take many months.[121] Phosphoramidites were chemically advantageous for application in solid-phase synthesis in relation to the preceding phosphorus III phosphite triester reagents. Phosphoramidite reagents are moderately stable under normal laboratory conditions in comparison to their chloride analogues and can be isolated and stored for use in tetrazole-catalyzed nucleotide couplings at room temperature without the formation of undesired pyridinium chloride precipitates or the use of hazardous tetrahydrofuran solvent and unstable phosphorodichloridite reagents.

Although 5' DMT is a conventional protecting group for phosphoramidite reagents, the TBS group was preferred for 5' protection due to the arguments outlined previously for the phosphite triester approach. Reports of the preparation of analogous 5' TBS phosphoramidites were unavailable in the literature. Thus, a series of silyl protected

phosphoramidite reagents of the canonical deoxynucleosides were prepared with standard

N-Ibu or *N*-benzoyl (Bz) base protection and 5' TBS ether protection using Singha *et al.'s*

original procedure for the 2-cyanoethyl,-*N,N*-diisopropylamino phosphoramidites

(Scheme 2.4).[112-114,122] Isolation of the moisture and air sensitive phosphoramidites **8a-d**

Scheme 2.4   Preparation of silyl protected phosphoramidite reagents.
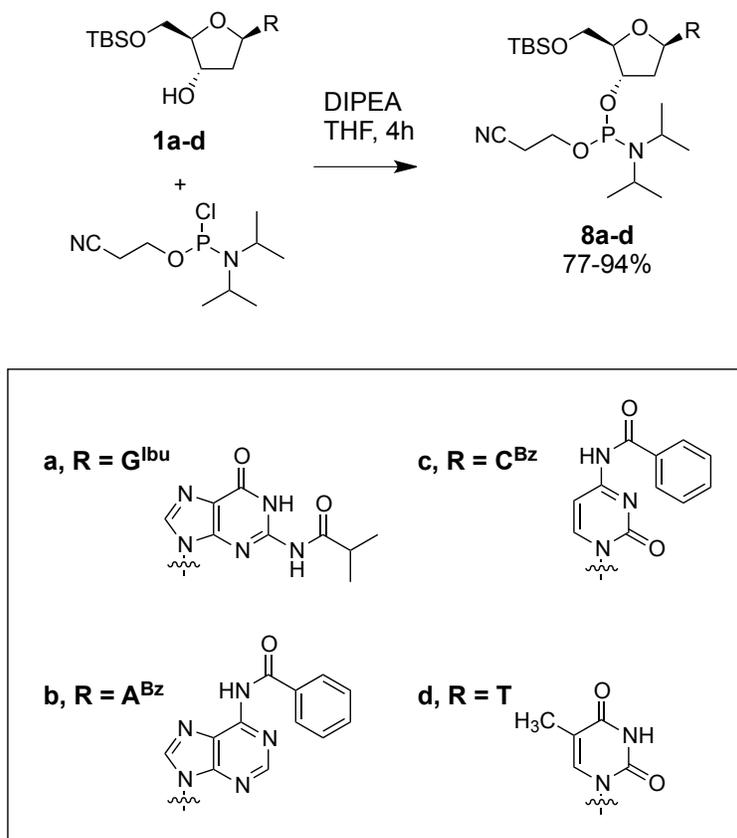
as a mixture of diastereomers was confirmed by the appearance of two singlets in the 150

ppm range in the ^{31}P NMR spectrum (Figure 2.3).

Couplings were performed at room temperature in tetrahydrofuran or dichloromethane with 16 equivalents of phosphoramidites **8a-d** and 24 equivalents of 5-(ethylthio)-tetrazole with foot-functionalized cavitand **2** or rim-functionalized cavitand **6** (Schemes 2.5-2.6). Simplified two step deblocking was carried out by simultaneous removal of the ammonia sensitive base and phosphate protecting groups followed by

Figure 2.3 ^{31}P NMR spectrum of silyl protected phosphoramidite **8d**.
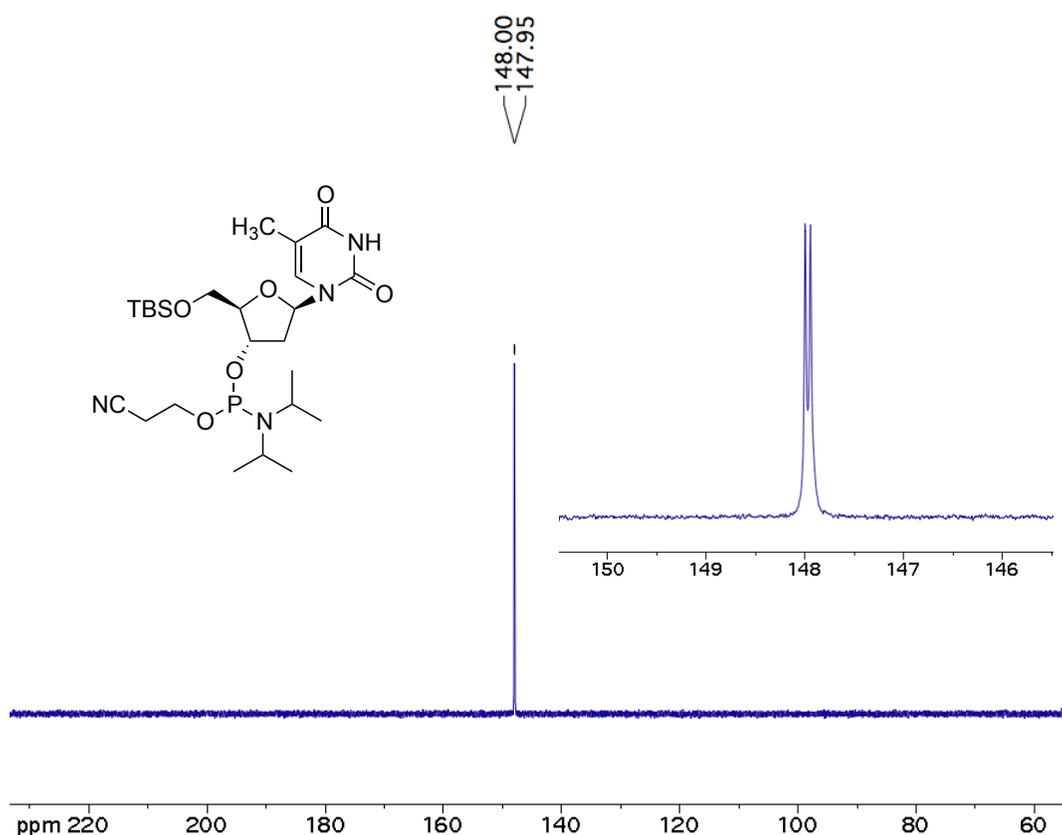
TBAF treatment and purification. The *n*-tetrabutylammonium salts **3a-d** and **7** were afforded in improved yield over the phosphite triester methodology.

A route to synthesizing base-protected analogues of cavitand-nucleotide conjugates **3a-d** was revealed (Scheme 2.7). Base-protected cavitand-nucleotide conjugates **9a-c** were desired for control studies based on analytical techniques such as circular dichroism (CD) and diffusion ordered NMR spectroscopy (DOSY). Probing of the self-assembly potential of the unblocked conjugates could be aided by comparison with an analogue conjugate with blocked hydrogen bonding capacity.

Scheme 2.5 Phosphoramidite synthesis of cavitand-nucleotide conjugates at the feet position.
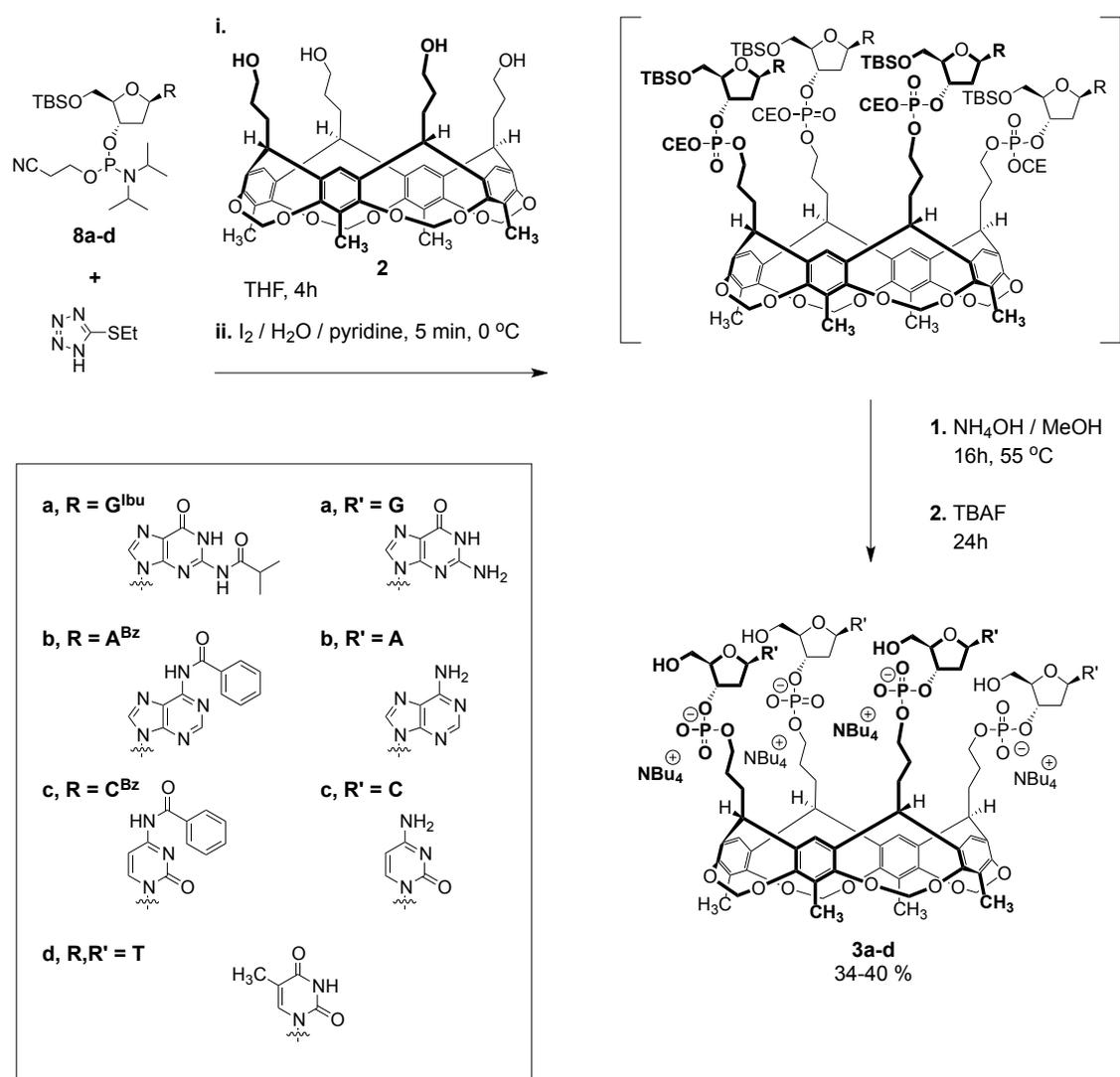
Selective removal of the 2-cyanoethyl (CE) phosphate protecting group was accomplished by substitution of a sterically hindered tertiary amine base in anhydrous pyridine in place of ammonia treatment.

Scheme 2.6 Phosphoramidite synthesis of guanine base cavitand-nucleotide conjugate at rim position.
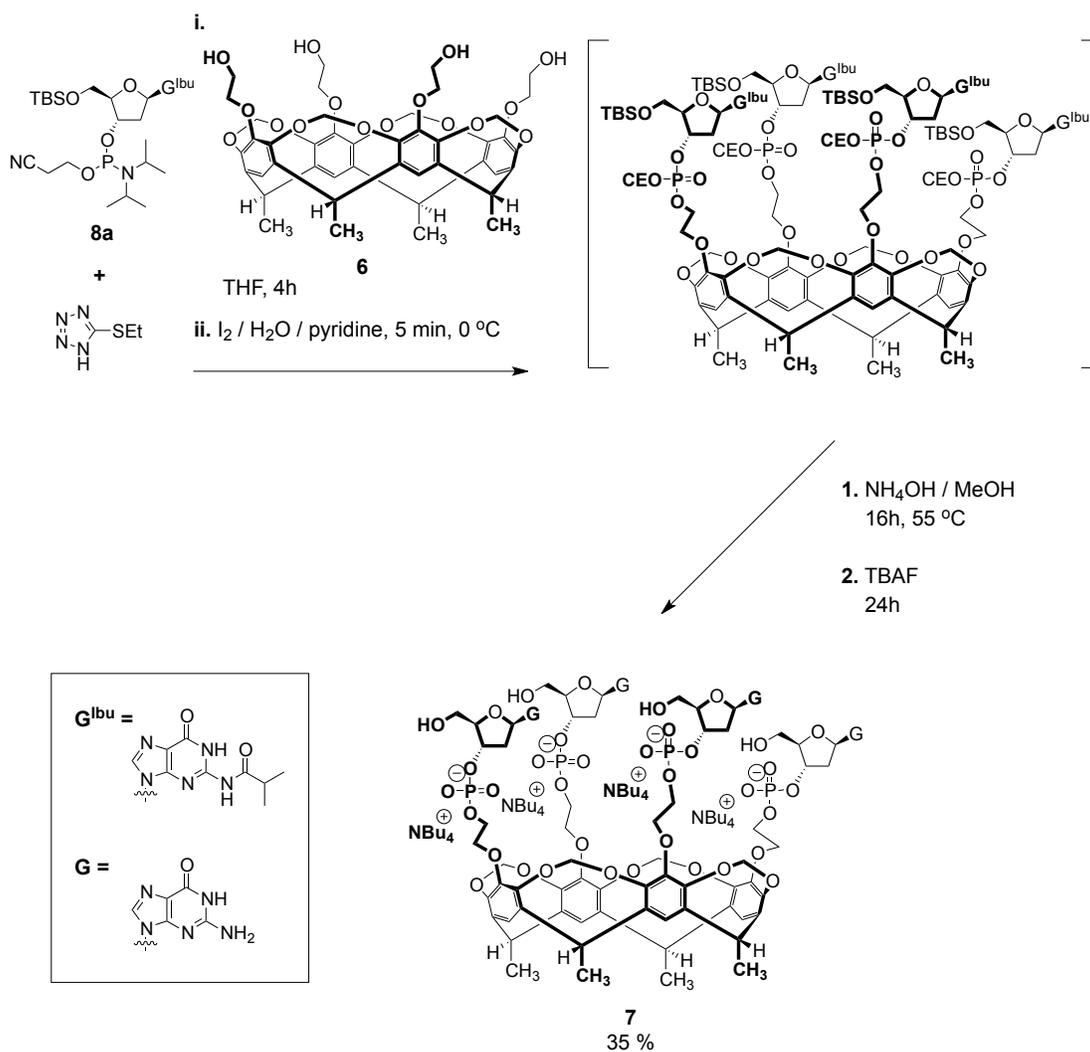
Scheme 2.7 Phosphoramidite synthesis of base-protected cavitand-nucleotide conjugate at the feet position.

a, R = G^Ibu

b, R = A^Bz

c, R = C^Bz

9a-c
18-42 %
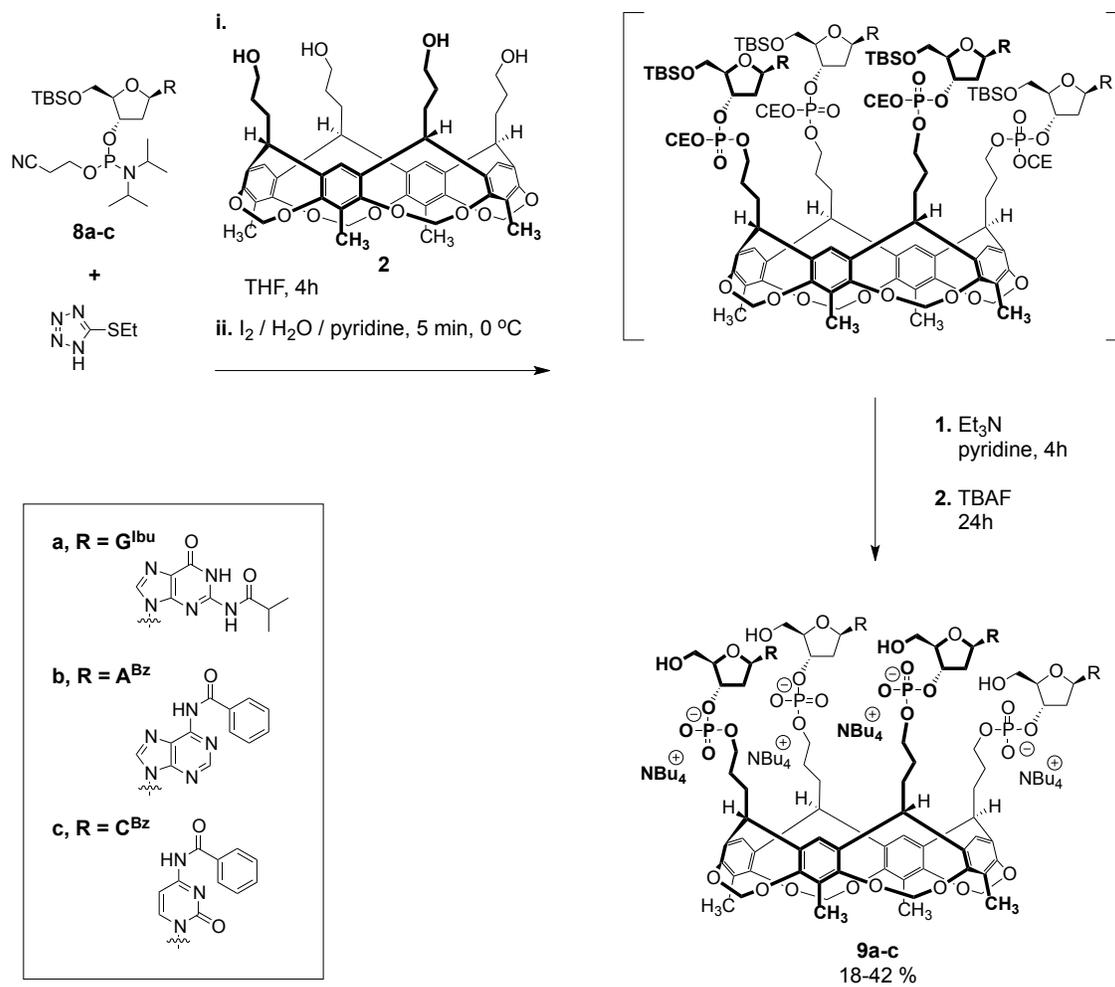
2.5 Experimental Procedures

General Procedures. All chemicals were purchased from standard suppliers. Reaction solvents were dried prior to use. Acetone, dichloromethane, ethyl acetate, and hexanes were HPLC grade and were dried over 3 Å molecular sieves. Dichloromethane, ethyl

acetate, and hexanes used as reaction solvents or during the work-up for the synthesis of phosphoramidites **8a-d** or for cavitand-nucleotide conjugate **3c** were bubbled with argon gas 30 min immediately prior to use. Dry tetrahydrofuran was obtained by distillation from sodium / benzophenone under a nitrogen atmosphere. Reactions were monitored by thin-layer chromatography on silica gel 60 F254nm aluminum sheets using methanol / dichloromethane developing solvent mixtures. Developed thin-layer chromatography plates were visualized at 254 nm followed by staining with aqueous 10 % H_2SO_4 solution and heating. Silica gel used for column chromatography was 230-400 mesh. All reported ^{1}H NMR and ^{13}C NMR chemical shifts were calibrated using the residual solvent signal relative to tetramethylsilane. Phosphorus NMR was performed with proton decoupling, and all reported ^{31}P NMR chemical shifts were referenced relative to aqueous 85 % H_3PO_4. The general numbering conventions used for NMR assignments are given below.
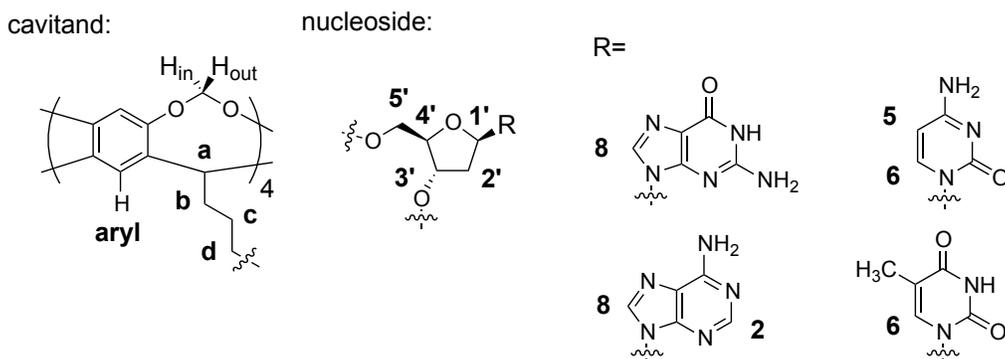
General Phosphite Triester Synthetic Procedure for Cavitand-Nucleotide Conjugates 3a and 7. To a stirred solution of 2,2,2-trichloroethylphosphorodichloridite

(0.13 mL, 0.83 mmol, 14 eq) and 2,6-lutidine (0.4 mL, 3.4 mmol, 57 eq) in 3 mL of tetrahydrofuran at -78 °C under an argon atmosphere was added dropwise over 5 min nucleoside **1a** (0.44 g, 0.97 mmol, 16 eq, coevaporated two times with 2 mL of tetrahydrofuran) in 2 mL of tetrahydrofuran. The reaction flask was occasionally removed from the cold bath and swirled while stirring was maintained for 1 h at -78 °C. The reaction mixture was warmed to 0 °C, and cavitand **2** or cavitand **6** (50 mg, 0.06 mmol, 1 eq, coevaporated two times with 2 mL of tetrahydrofuran) was added dropwise over 5 mins in 4 mL of tetrahydrofuran. Stirring was maintained for 4 h at 0 °C. A solution of 0.1 M iodine in tetrahydrofuran / pyridine / water (14 mL, 80:20:2) was added, and stirring was continued for 5 min at 0 °C. The reaction solvent was removed by rotary evaporation, and the residue was dissolved in 40 mL of dichloromethane and washed with 100 mL of 5 % aqueous sodium metabisulfite. The aqueous layer was extracted two more times with 40 mL of dichloromethane. The combined organic layer was washed with 100 mL of brine and dried over anhydrous $MgSO_4$. The solvent was removed by rotary evaporation, and the crude product was purified by flash chromatography on silica gel using methanol / dichloromethane as a solvent system (1:20 then 1:10). Complete deblocking was performed without isolation of the partially deblocked intermediates. The fully blocked product mixture was coevaporated two times in 2 mL of dichloromethane. Acetylacetone (0.2 mL) and Zn / Cu couple (two times mass of the fully blocked product mixture) were added, and the reaction mixture was suspended in 20 mL of pyridine. The reaction was stirred for 4 h under an argon atmosphere. A 40 mL solution of 1:1 pyridine / water and 4 g of DOWEX 50WX8 50-100 (NH_4^+ form) ion-exchange resin were added, and stirring was continued for 15 mins.

The resin was filtered off, and the solvent was removed by rotary evaporation. The residue was dissolved in 6 mL of methanol and 6 mL of concentrated aqueous ammonia and stirred for 16 h at 55 °C. After evaporation of the solvent, the residue was dissolved in 3 mL of 1 M TBAF in tetrahydrofuran and stirred for 24 h. The reaction was diluted in 15 mL of water, loaded onto a Sep-Pak Vac 12 cc (2 g) C18 cartridge, and purified by step-gradient elution with pure water to pure methanol in 10 % increments with 40 mL per step. Fractions were monitored by UV / VIS spectrophotometry at 260 nm, and products were eluted at approximately 50 % methanol composition. Lyophilization from water afforded compounds **3a** and **7** as white powders in n-tetrabutylammonium salt form.

3a. 30 mg (16 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 7.90 (s, 4H, H-8), 7.10 (s, 4H, aryl), 6.21 (t, J = 7.0 Hz, 4H, CH-1'), 5.68 (d, J = 7.0 Hz, 4H, H$_{out}$), 5.03 (br. s, 4H, CH-3'), 4.74 (t, J = 8.0 Hz, 4H, CH-a), 4.21 (br, 4H, CH-4'), 4.09 (d, J = 7.0 Hz, 4H, H$_{in}$), 4.00 (m, 8H, CH$_2$-d), 3.82 (m, 8H, CH$_2$-5'), 2.66 (m, 4H, CH$_a$-2'), 2.42 (m, 12H, CH$_2$-b, CH$_b$-2'), 1.76 (s, 12H, methyl), 1.80 (overlap, 8H, CH$_2$-c); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 159.4, 155.1, 154.7, 152.3, 138.6, 138.4, 125.3, 119.2, 118.4, 99.8, 88.4, 86.1, 76.9, 65.7, 63.8, 59.5, 40.1, 37.0, 29.7, 26.2, 24.8, 20.7, 14.0, 10.0; ^{31}P-NMR (162 MHz, CD$_3$OD) δ 2.3 (s); MS (MALDI-TOF) calculated for C$_{88}$H$_{103}$N$_{20}$O$_{36}$P$_4$ (M-H)$^-$: 2139.6; found: 2142.0.

7. 49 mg (26 % overall yield); ^{1}H-NMR (400 MHz, D$_2$O) δ 8.03 (s, 4H, H-8), 6.99 (s, 4H, aryl), 6.25 (t, J = 7.0 Hz, 4H, CH-1'), 5.88 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.94 (br, 4H, CH-3'),

4.32 (q, J = 7.0 Hz, 4H, CH-a), 4.23 (br, 4H, CH-4'), 4.2-3.8 (m, 20H, H$_{in}$, -OCH$_2$CH$_2$O-), 3.75 (m, 8H, CH$_2$-5'), 2.88 (m, 4H, CH$_a$-2'), 2.76 (m, 4H, CH$_b$-2'), 1.60 (overlap, 12H, methyl); ^{13}C NMR (101 MHz, D$_2$O) δ 158.4, 153.9, 151.2, 147.3, 142.8, 139.9, 137.5, 115.8, 115.1, 100.3, 86.4, 83.9, 76.3, 73.7, 65.1, 61.9, 58.2, 37.2, 31.2, 23.3, 19.3, 15.2, 12.8; ^{31}P-NMR (162 MHz, D$_2$O) δ -2.4 (s); MS (MALDI-TOF) calculated for C$_{84}$H$_{95}$N$_{20}$O$_{40}$P$_4$ (M-H)$^-$: 2147.5; found: 2149.2.

Cavitand 5. To a stirred solution of cavitand **4** (1.95g, 2.97 mmol, 1 eq, coevaporated two times with 10 mL of acetone) in 100 mL of acetone under an argon atmosphere was added anhydrous potassium carbonate (2.50 g, 18.1 mmol, 6 eq) and ethyl bromoacetate (2.8 mL, 25.1 mmol, 8.5 eq). Stirring was maintained under reflux for 4 days. After removal of the solvent by rotary evaporation, the residue was partioned between 100 mL of dichloromethane and 200 mL of 2 % HCl solution extracted three times. The combined organic layer was washed with 100 mL of brine, dried over anhydrous MgSO$_4$, and the solvent was removed by rotary evaporation. The residue was dissolved in 15 mL of dichloromethane, precipitated into stirring 200 mL of hexanes, suction filtered, and dried under high vacuum to give 2.34 g (78 % yield) of compound **5** as a white solid.

^{1}H NMR (400 MHz, CDCl$_3$) δ 6.93 (s, 4H, aryl), δ 5.75 (d, J = 7.5 Hz, 4H, H$_{out}$), δ 4.93 (q, J = 7.5 Hz, 4H, CH-a), δ 4.52 (s, 8H, -C$\underline{H}_2$CO-, 8H), δ 4.44 (d, J = 7.5 Hz, 4H, H$_{in}$), δ 4.23 (q, J = 7.0 Hz, 8H, -OC$\underline{H}_2$CH$_3$), δ 1.71 (d, J = 7.5 Hz, 12H, methyl), δ 1.29 (t, J = 7.0 Hz, 12H, -OCH$_2$C$\underline{H}_3$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 169.4, 147.1, 143.9, 139.8,

113.9, 100.0, 70.2, 61.1, 31.2, 16.0, 14.3; HRMS (ESI) calculated for $C_{52}H_{56}O_{20}Na$ $(M+Na)^+$: 1023.3263; found: 1023.3248.

Cavitand 6. To a stirred solution of cavitand **5** (1.00 g, 1.00 mmol, 1 eq, coevaporated two times in 5 mL of tetrahydrofuran) in 25 mL of tetrahydrofuran under an argon atmosphere and cooled to -78 °C was added $LiAlH_4$ (0.50 g, 13 mmol, 13 eq). The reaction mixture was allowed to warm to room temperature over 2 hours. The reaction mixture was cooled to 0 °C, quenched with 10 mL of ethyl acetate followed by 10 mL of water, and the solvent was reduced by rotary evaporation. The residue was partitioned between 75 mL of ethyl acetate and 200 mL of 3 M H_2SO_4 and extracted three times. The combined organic layer was washed with 100 mL of water, 100 mL of saturated aqueous $NaHCO_3$, and 100 mL of brine, dried over anhydrous $MgSO_4$, and the solvent was removed by rotary evaporation. The crude product was purified by flash chromatography on silica gel using ethanol / dichloromethane as a solvent system (1:20 then 1:10) and dried under high vacuum to give 0.48 g (58 % yield) of compound **6** as white solid.

^1H NMR (400 MHz, CDCl$_3$) δ 6.96 (s, 4H, aryl), δ 5.88 (d, J = 7.0 Hz, 4H, H$_{out}$), δ 4.91 (q, J = 7.0 Hz, 4H, CH-a), δ 4.44 (d, J = 7.0 Hz, 4H, H$_{in}$), δ 4.11 (br, 8H, -OC$\underline{H}_2$CH$_2$OH), δ 3.64 (br, 8H, -OCH$_2$C$\underline{H}_2$OH), δ 3.14 (br. s, 4H, OH), δ 1.74 (d, J = 7.0 Hz, 12H, methyl); ^{13}C NMR (101 MHz, CDCl$_3$) δ 147.8, 143.5, 140.0, 114.1, 99.7, 75.6, 61.5, 31.3, 15.9; HRMS (ESI) calculated for $C_{44}H_{48}O_{16}Na$ $(M+Na)^+$: 855.2840; found: 855.2829.

General Synthetic Procedure for Phosphoramidites 8a-d. To a stirred solution of nucleoside **1a-d** (3.30 mmol, 1 eq, coevaporated three times in 2 mL of tetrahydrofuran) and *N,N*-diisopropylethylamine (2.3 mL, 13 mmol, 4 eq) in 15 mL of tetrahydrofuran under an argon atmosphere was added dropwise over 2 mins 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (0.90 mL, 4 mmol, 1.2 eq). Stirring was maintained for 4 h. After removal of the solvent by rotary evaporation, the residue was dissolved in 50 mL of ethyl acetate and washed with 50 mL of saturated aqueous $NaHCO_3$, 50 mL of brine, and dried over anhydrous $MgSO_4$. The solvent was removed by rotary evaporation, and the residue was dissolved in 3 mL of ethyl acetate (dichloromethane was used in place of ethyl acetate to dissolve **8c**). The solution was added dropwise into 400 mL of stirring hexanes at -78 ºC under argon, and the resultant precipitate was suction filtered, and coevaporated three times in tetrahydrofuran (dichloromethane was used in place of tetrahydrofuran for **8c**), and dried under high vacuum. The white crystalline compounds were stable over months when stored under argon at -20 ºC.

8a. 2.03 g (94 % yield); mixture of diastereomers; ^{31}P-NMR (162 MHz, CDCl$_3$) δ 150.2 (s), 148.7 (s); HRMS (ESI) calculated for $C_{29}H_{51}N_7O_6PSi$ (M+H)$^+$: 652.3408; found: 652.3417.

8b. 1.81 g (88 % yield); mixture of diastereomers; ^{31}P-NMR (162 MHz, CDCl$_3$) δ 149.9 (s), 149.8 (s); HRMS (ESI) calculated for $C_{32}H_{49}N_7O_5PSi$ (M+H)$^+$: 670.3302; found: 670.3317.

8c. 1.51 g (77 % yield); mixture of diastereomers;^{31}P-NMR (162 MHz, CDCl$_3$) δ 148.0 (s), 147.95 (s); HRMS (ESI) calculated for C$_{31}$H$_{49}$N$_5$O$_6$PSi (M+H)$^+$: 646.3190; found: 646.3179.

8d. 1.60 g (95 % yield); mixture of diastereomers;^{31}P-NMR (162 MHz, CDCl$_3$) δ 149.4 (s), 149.4 (s); HRMS (ESI) calculated for C$_{25}$H$_{46}$N$_4$O$_6$PSi (M+H)$^+$: 557.2924; found: 557.2923.

General Phosphoramidite Synthetic Procedure for Cavitand-Nucleotide Conjugates 3a-d and 7. To a stirred solution of phosphoramidite **8a-d** (0.97 mmol, 16 eq) and 5-(ethylthio)tetrazole (0.19 g, 1.5 mmol, 24 eq, coevaporated two times in 2 mL of tetrahydrofuran) in 6 mL of tetrahydrofuran under an argon atmosphere was added dropwise over 5 mins cavitand **2** or cavitand **6** (50 mg, 0.60 mmol, 1 eq, coevaporated two times in 2 mL of tetrahydrofuran) in 4 mL of tetrahydrofuran (dichloromethane was used in place of tetrahydrofuran for the reaction of phosphoramidite **8c**). Stirring was maintained for 3 h. The reaction mixture was cooled to 0°C, 0.2 mL of water was added, and stirring was continued for 1 min. A solution of 0.1 M iodine in tetrahydrofuran / pyridine / water (14 ml, 80:20:2) was added, and stirring was continued for 5 mins at 0 °C. The reaction solvent was removed by rotary evaporation, and the residue was dissolved in 40 mL of dichloromethane and washed with 100 mL of 5 % aqueous sodium metabisulfite. The aqueous layer was extracted two more times with 40 mL of dichloromethane. The combined organic layer was washed with 100 mL of saturated aqueous NaHCO$_3$, 100 mL of brine, and dried over anhydrous MgSO$_4$. The solvent was

removed by rotary evaporation, and the crude product was immediately purified by flash chromatography on silica gel using methanol / dichloromethane as a solvent system (1:20 then 1:10). Deblocking was performed by dissolving the fully blocked product mixture in 6 mL of methanol and 6 mL of concentrated aqueous ammonia followed by stirring for 24 h at 55 °C. After evaporation of the solvent, the residue was dissolved in 3 mL of 1 M TBAF in tetrahydrofuran and stirred for 24 h at room temperature. The reaction was diluted in 15 mL of water, loaded onto a Sep-Pak Vac 12 cc (2 g) C18 cartridge, and purified by step-gradient elution with pure water to pure methanol in 10 % increments with 40 mL per step. Fractions were monitored by UV / VIS spectrophotometry at 260 nm, and products were eluted at approximately 50 % methanol composition. Lyophilization from water afforded final compounds **3a-d** as white powders in *n*-tetrabutylammonium salt form.

3a. 67 mg (36 % overall yield)

3b. 74 mg (40 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.36 (s, 4H, H-2), 8.19 (s, 4H, H-8), 7.16 (s, 4H, aryl), 6.44 (t, J = 6.0 Hz, 4H, CH-1'), 5.67 (d, J = 7.0 Hz, 4H, H$_{out}$), 5.00 (br, 4H, CH-3'), 4.79 (t, J = 8.0 Hz, 4H, CH-a), 4.29 (br, 4H, CH-4'), 4.11 (d, J = 7.0 Hz, 4H, H$_{in}$), 4.01 (m, 8H, CH$_2$-d), 3.80 (br, 8H, CH$_2$-5'), 2.80 (m, 4H, CH$_a$-2'), 2.57 (m, 4H, CH$_b$-2'), 2.44 (m, 8H, CH$_2$-b), 1.77 (s, 12H, methyl), 1.80 (overlap, 8H, CH$_2$-c); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 156.3, 154.6, 152.0, 149.7, 142.1, 138.5, 125.1, 120.7, 119.1, 99.9, 88.9, 86.9, 77.1, 65.9, 63.6, 59.5, 40.8, 36.8, 29.8, 27.0, 24.8, 20.7,

14.0, 10.4; ^{31}P-NMR (162 MHz, CD$_3$OD) δ 2.1 (s); MS (MALDI-TOF) calculated for C$_{88}$H$_{103}$N$_{20}$O$_{32}$P$_4$ (M-H)$^-$: 2075.6; found: 2076.5.

3c. 54 mg (34 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.10 (d, J = 7.0 Hz, 4H, H-6), 7.15 (s, 4H, aryl), 6.30 (t, J = 6.0 Hz, CH-1'), 6.08 (d, J = 7.0 Hz, 4H, H-5), 5.78 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.78 (br, 8H, CH-3', CH-a), 4.16 (br., 8H, CH-4', H$_{in}$), 3.96 (br, 8H, CH$_2$-d), 3.80 (br, 8H, CH$_2$-5'), 2.47 (br, 4H, CH$_a$-2'), 2.34 (br, 8H, CH$_2$-b), 2.06 (br, 4H, CH$_b$-2'), 1.89 (s, 12H, methyl), 1.72 (overlap, 8H, CH$_2$-c); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 166.1, 154.7, 143.4, 139.0, 138.4, 125.8, 119.1, 99.9, 96.3, 88.1, 87.9, 86.8, 76.0, 65.5, 63.0, 59.5, 40.8, 36.7, 29.8, 26.5, 24.8, 20.7, 13.9, 10.6; ^{31}P-NMR (162 MHz, D$_2$O) δ -2.8 (s); MS (MALDI-TOF) calculated for C$_{84}$H$_{103}$N$_{12}$O$_{36}$P$_4$ (M-H)$^-$: 1979.6; found: 1982.1.

3d. 64 mg (35 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 7.87 (s, 4H, H-6), 7.19 (s, 4H, aryl), 6.30 (t, J = 7.0 Hz, 4H, CH-1'), 5.79 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.7-4.9 (overlap, 8H, CH-3',CH-a), 4.21 (d, J = 7.0, 4H, H$_{in}$), 4.14 (br, 4H, CH-4'), 3.99 (br. 8H, CH$_2$-d), 3.78 (br, 8H, CH$_2$-5'), 2.44 (br, 8H, CH$_2$-b), 2.1-2.3 (m, 8H, CH$_2$-2'), 1.93 (s, 12H, methyl), 1.91 (s, 12H, methyl), 1.75 (overlap, 8H, CH$_2$-c); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 166.4, 154.8, 152.2, 138.6, 138.3, 125.4, 119.2, 111.6, 99.9, 87.9, 86.3, 76.5, 65.7, 62.9, 59.5, 40.4, 37.3, 29.9, 27.0, 24.8, 20.7, 14.0, 12.7, 10.3; ^{31}P-NMR (162 MHz, D$_2$O) δ -2.6 (s); MS (MALDI-TOF) calculated for C$_{88}$H$_{109}$N$_8$O$_{40}$P$_4$ (M+H)$^+$: 2041.6; found: 2043.9.

7. 65 mg (35 % overall yield)

General Phosphoramidite Synthetic Procedure for Base-Protected Cavitand-Nucleotide Conjugates 9a-c. The general phosphoramidite synthetic procedure for cavitand-nucleotide conjugates **3a-d** was performed with the following modifications. Treatment of the fully blocked product mixture with concentrated aqueous ammonia and methanol solution was avoided. To a stirred solution of the fully blocked product mixture coevaporated three times in 2 mL of pyridine and dissolved in 10 mL of pyridine was added 6 mL of triethylamine under an argon atmosphere. Stirring was maintained for 4 h. The solvent was removed by rotary evaporation, and the residue was coevaporated three times in pyridine. Subsequent TBAF treatment and purification was carried out as described previously to afford compounds **9a-c** as white powders in *n*-tetrabutylammonium salt form.

9a. 70 mg (34 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.26 (s, 4H, H-8), 7.09 (s, 4H, aryl), 6.32 (t, J = 7.0 Hz, 4H, CH-1'), 5.61 (d, J = 7.0 Hz, 4H, H$_{out}$), 5.01 (br, 4H, CH-3'), 4.68 (t, J = 8.0 Hz, 4H, CH-a), 4.19 (br, 4H, CH-4'), 4.06 (m, 8H, H$_{in}$, CH$_2$-d), 3.78 (br, 8H, CH$_2$-5'), 2.82 (m, 4H, -C$\underline{H}$(CH$_3$)$_2$) 2.67 (m, 4H, CH$_a$-2'), 2.57 (m, 4H, CH$_b$-2'), 2.37 (m, 12H, CH$_2$-b), 1.77 (s, 12H, methyl), 1.72 (overlap, 8H, CH$_2$-c), 1.21 (m, 24H, -CH(C$\underline{H}_3$)$_2$); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 182.2, 157.5, 154.8, 150.2, 149.7, 140.1, 138.6, 125.4, 121.4, 119.2, 100.0, 88.4, 85.5, 76.9, 66.3, 63.3, 59.5, 40.7, 37.6, 36.8, 30.2, 26.8, 24.8, 20.7, 19.5, 14.0, 10.2; ^{31}P-NMR (162 MHz, CD$_3$OD) δ 2.0 (s); MS (MALDI-TOF) calculated for C$_{104}$H$_{129}$N$_{20}$O$_{40}$P$_4$ (M+H)$^+$: 2421.7; found: 2424.5.

9b. 89 mg (42 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.69 (s, 4H, H-2), 8.65 (s, 4H, H-8), 8.08 (m, 8H, benzoyl-*o*), 7.61 (m, 4H, benzoyl-*p*), 7.53 (m, 8H, benzoyl-*m*), 7.12 (s, 4H, aryl), 6.54 (t, *J* = 7.0 Hz, 4H, CH-1'), 5.63 (d, *J* = 7.0 Hz, 4H, H$_{out}$), 5.02 (br, 4H, CH-3'), 4.73 (t, *J* = 8.0 Hz, 4H, CH-a), 4.29 (br. s, 4H, CH-4'), 4.05 (m, 8H, H$_{in}$, CH$_2$-d), 3.81 (br, 8H, CH$_2$-5'), 2.80 (m, 4H, CH$_a$-2'), 2.59 (m, 4H, CH$_b$-2'), 2.41 (m, 8H, CH$_2$-b), 1.74 (s, 12H, methyl), 1.74 (overlap, 8H, CH$_2$-c); ^{31}P-NMR (162 MHz, CD$_3$OD) δ 2.2 (s); MS (MALDI-TOF) calculated for C$_{116}$H$_{119}$N$_{20}$O$_{36}$P$_4$ (M-H)$^-$: 2491.7; found: 2494.8.

9c. 37 mg (18 % overall yield); ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.55 (d, *J* = 7.5 Hz, 4H, H-6), 7.97 (m, 8H, benzoyl-*o*), 7.40-7.70 (m, 16H, benzoyl-*m/p*. H-5), 7.19 (m, 4H, aryl), 6.30 (t, *J* = 6.5 Hz, 4H, CH-1'), 5.77 (d, *J* = 7.0 Hz, 4H, H$_{out}$), 4.87 (overlap, 4H, CH-3') 4.80 (overlap, 4H, CH-a), 4.27 (br, 4H, CH-4'), 4.19 (br, 4H, H$_{in}$), 4.03 (m, 4H, CH$_2$-d), 3.84 (br, 8H, CH$_2$-5'), 2.65 (m, 4H, CH$_a$-2'), 2.45 (br, 8H, CH$_2$-b), 2.24 (m, 4H, CH$_b$-2'), 1.90 (s, 12H, methyl), 1.6-1.8 (overlap, 8H, CH$_2$-c); ^{13}C-NMR (101 MHz, CD$_3$OD) δ 169.0, 164.6, 157.6, 154.9, 146.9, 138.7, 134.8, 134.1, 129.8, 129.3, 125.5, 119.3, 100.1, 98.5, 88.7, 88.5, 76.5, 66.2, 62.7, 59.5, 53.7, 41.7, 37.3, 30.0, 26.7, 24.8, 20.7, 14.0, 10.3; ^{31}P-NMR (162 MHz, CD$_3$OD) δ 0.6 (s); MS (MALDI-TOF) calculated for C$_{112}$H$_{121}$N$_{12}$O$_{40}$P$_4$ (M+H)$^+$: 2397.7; found: 2400.0.

2.6 Supplementary ^{1}H ^{1}H COSY NMR Spectra

Supplementary ^{1}H ^{1}H COSY NMR spectra begins on the following page.

Figure 2.4   ¹H ¹H COSY NMR Spectrum of **3a** in CD₃OD at 400 MHz.
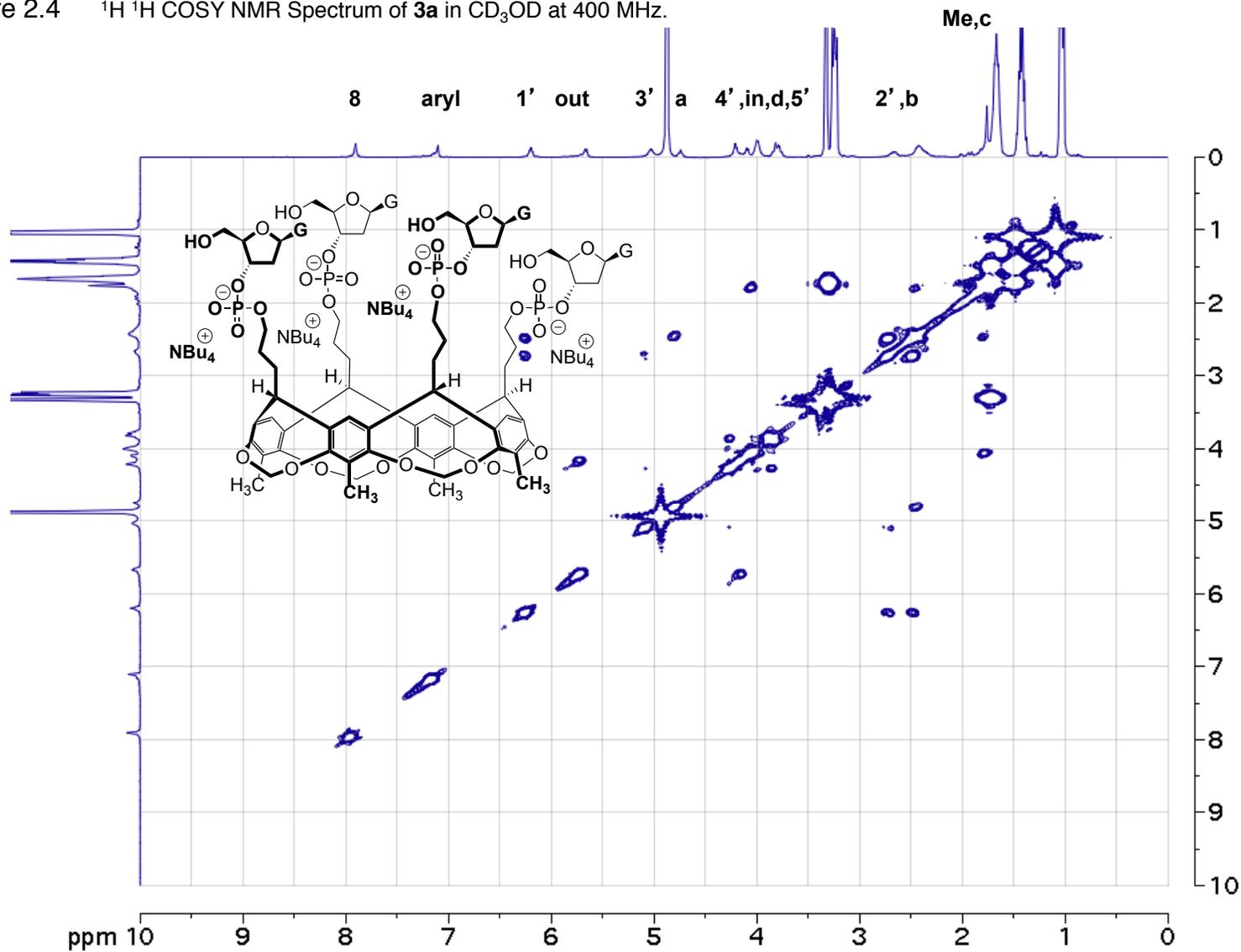
Figure 2.5 ^{1}H ^{1}H COSY NMR Spectrum of **3b** in CD$_3$OD at 400 MHz.
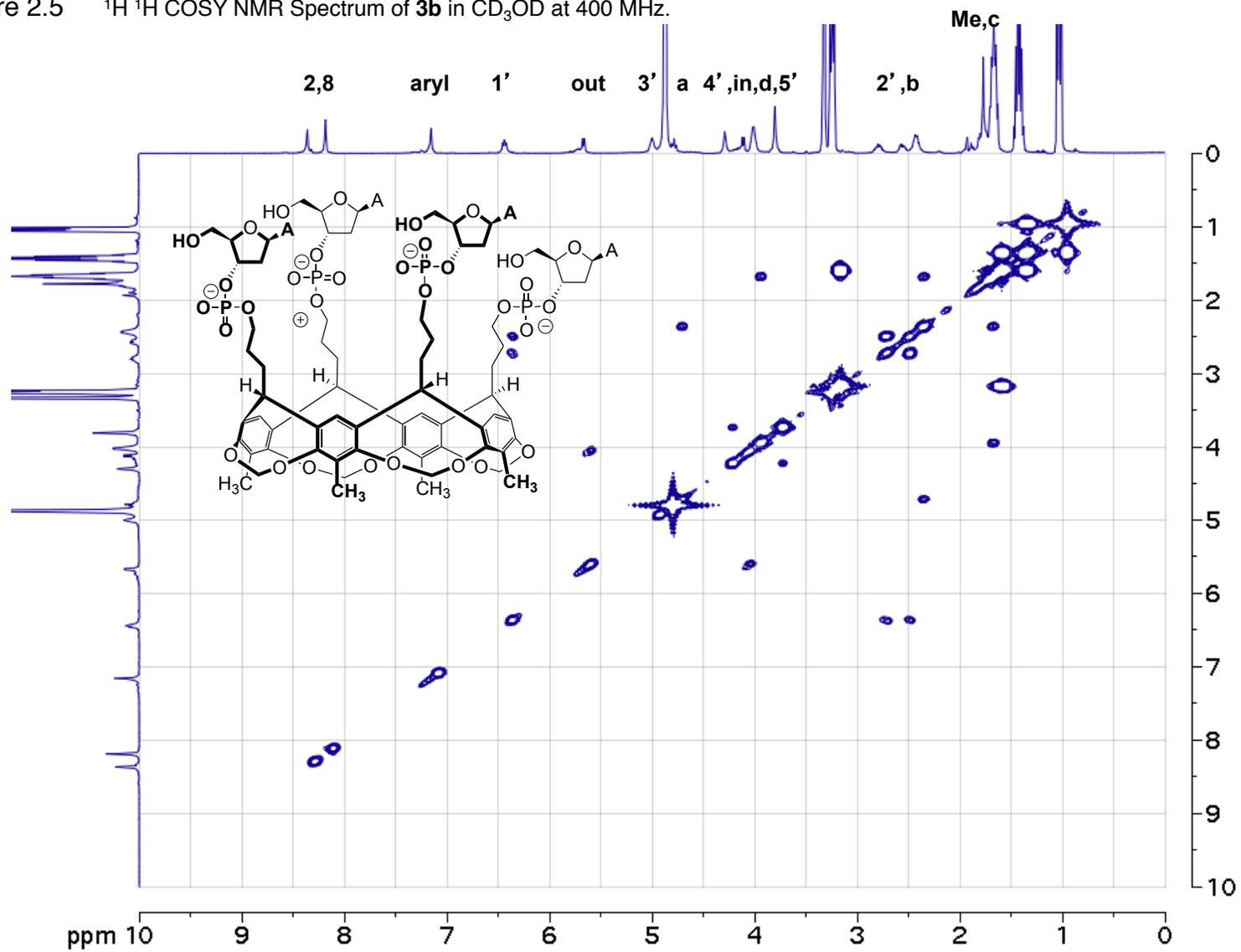
Figure 2.6   ¹H ¹H COSY NMR Spectrum of **3c** in CD₃OD at 400 MHz.
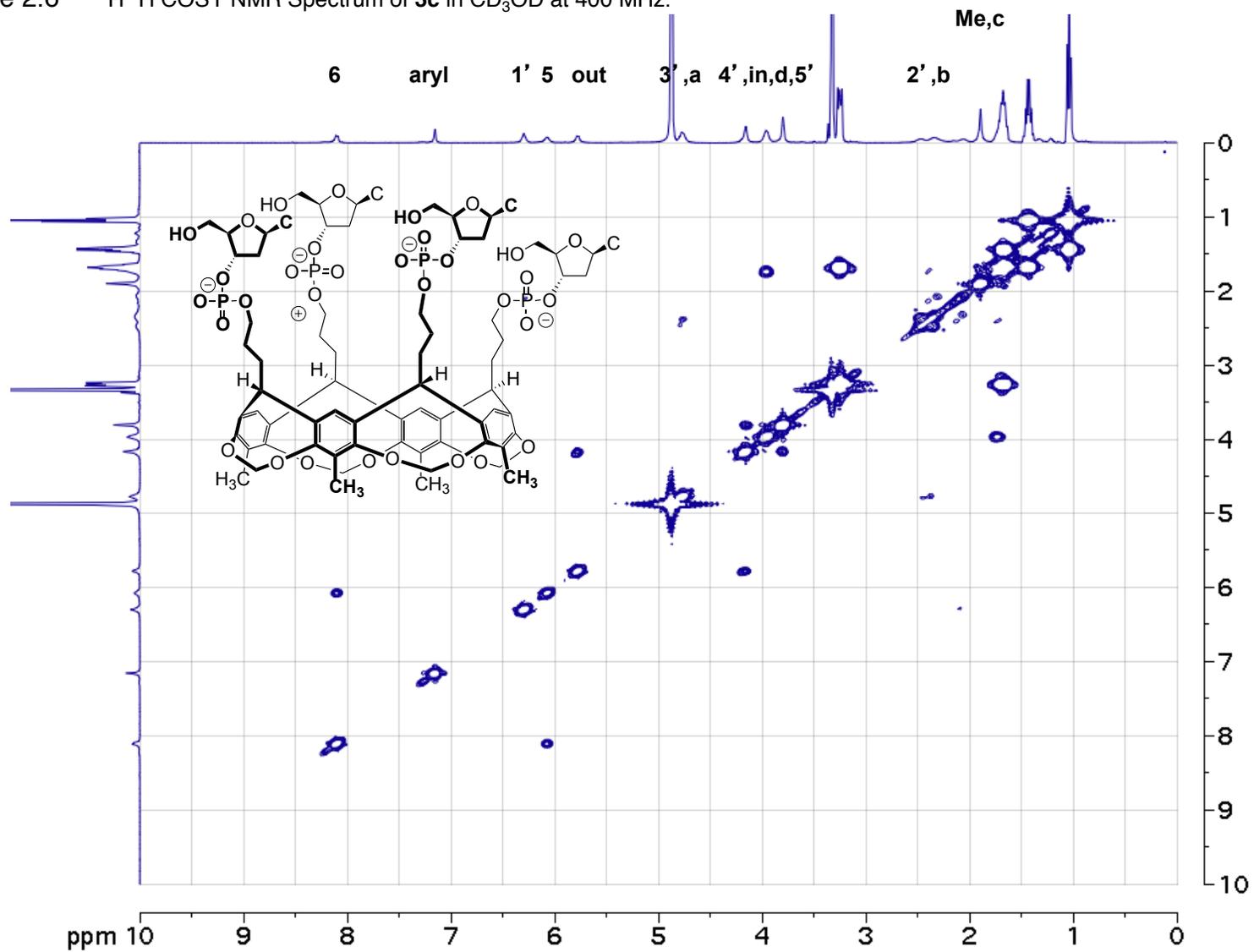
Figure 2.7 ¹H ¹H COSY NMR Spectrum of **3d** in CD₃OD at 400 MHz.
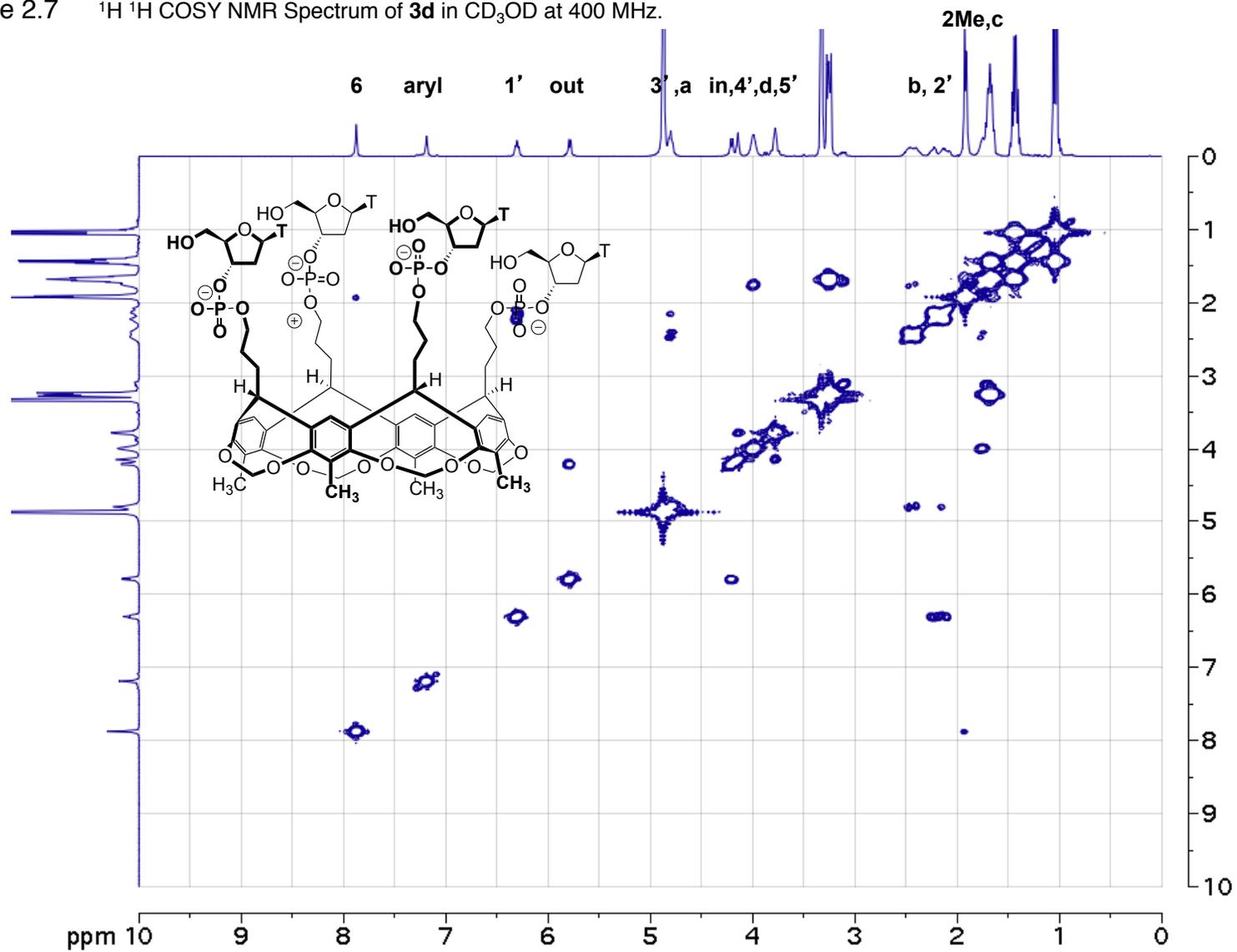
Figure 2.8   ¹H ¹H COSY NMR Spectrum of **7** in D₂O at 400 MHz.

69
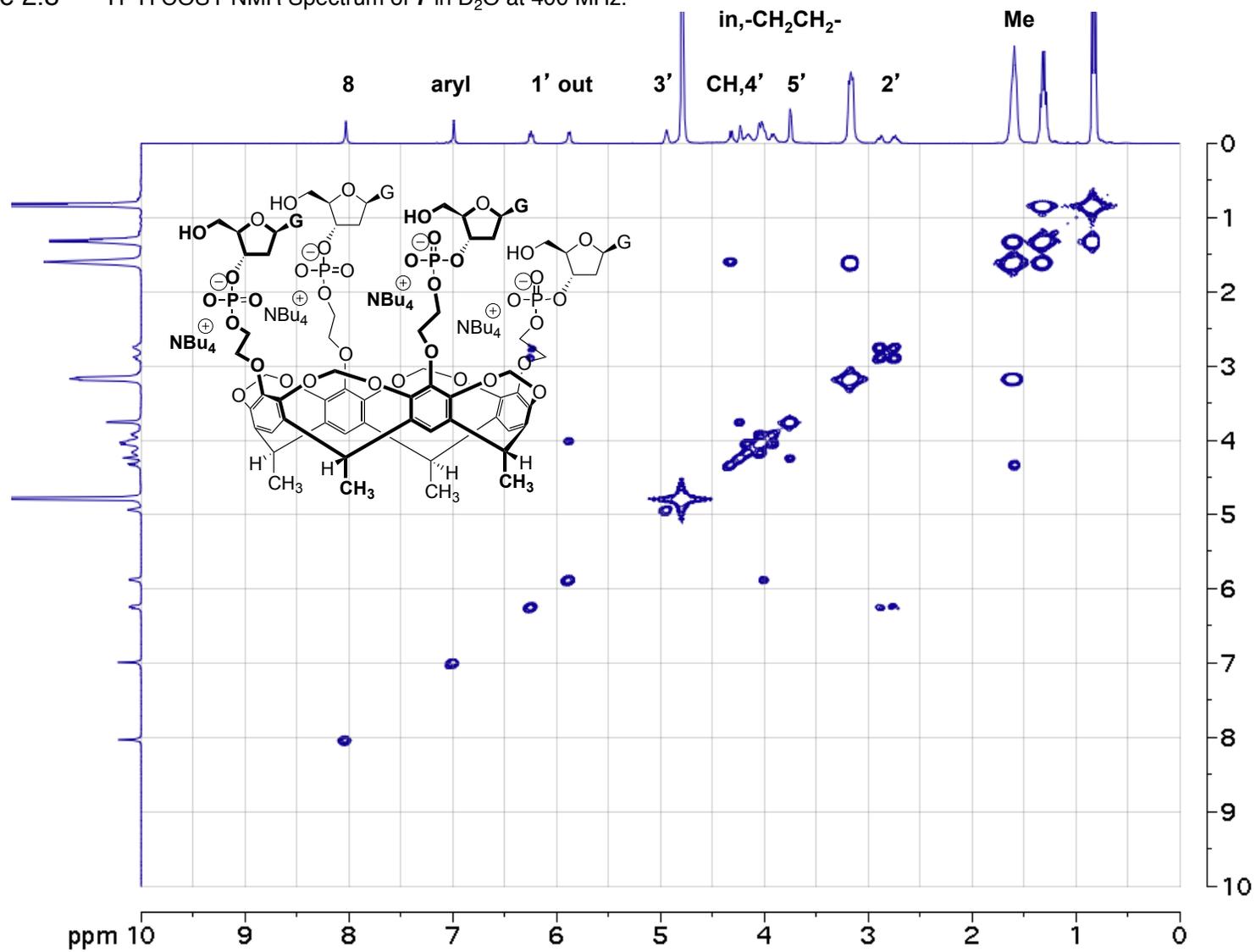
Figure 2.9 ^1H ^1H COSY NMR Spectrum of **9a** in CD_3OD at 400 MHz.
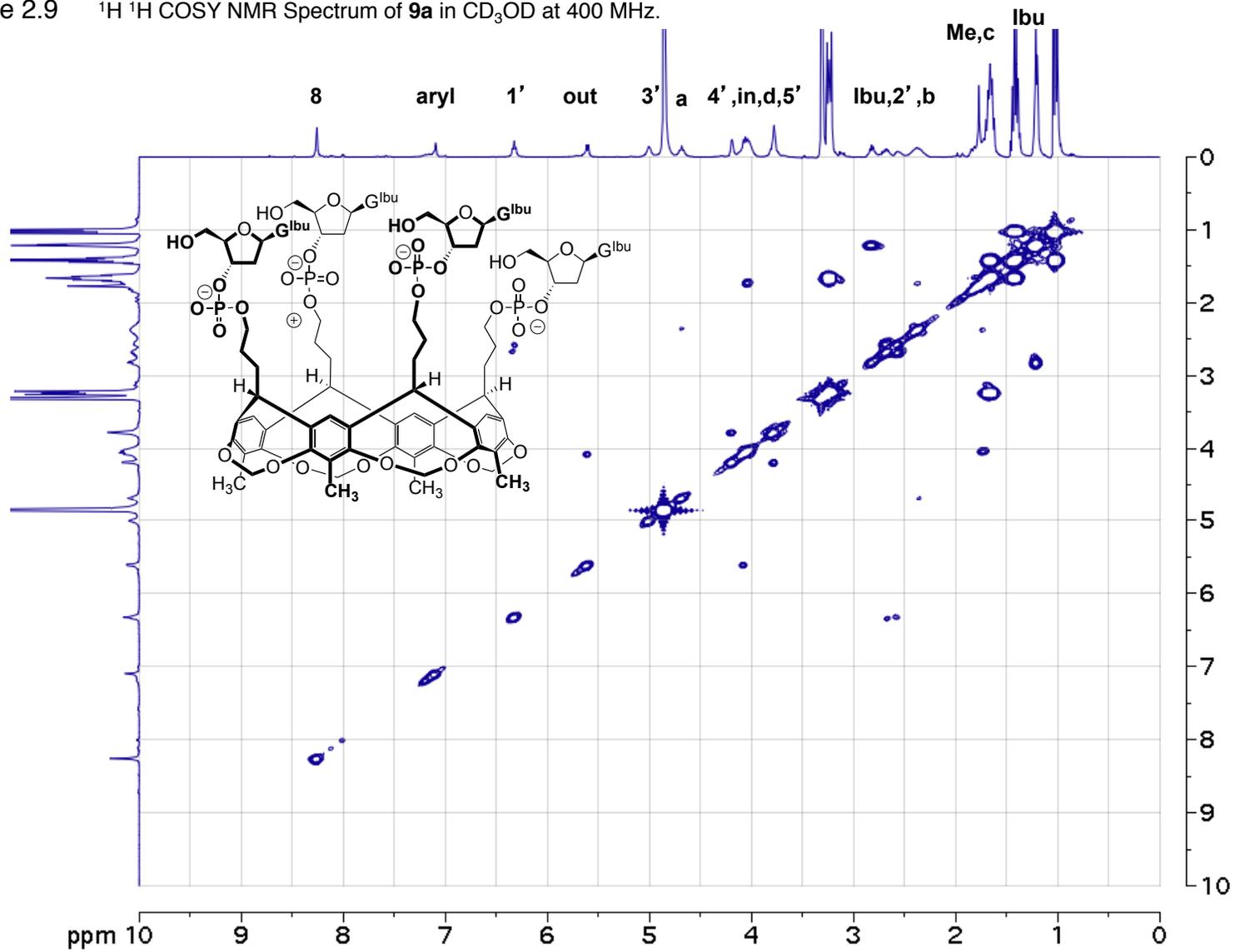
Figure 2.10 ¹H ¹H COSY NMR Spectrum of **9b** in CD$_3$OD at 400 MHz.
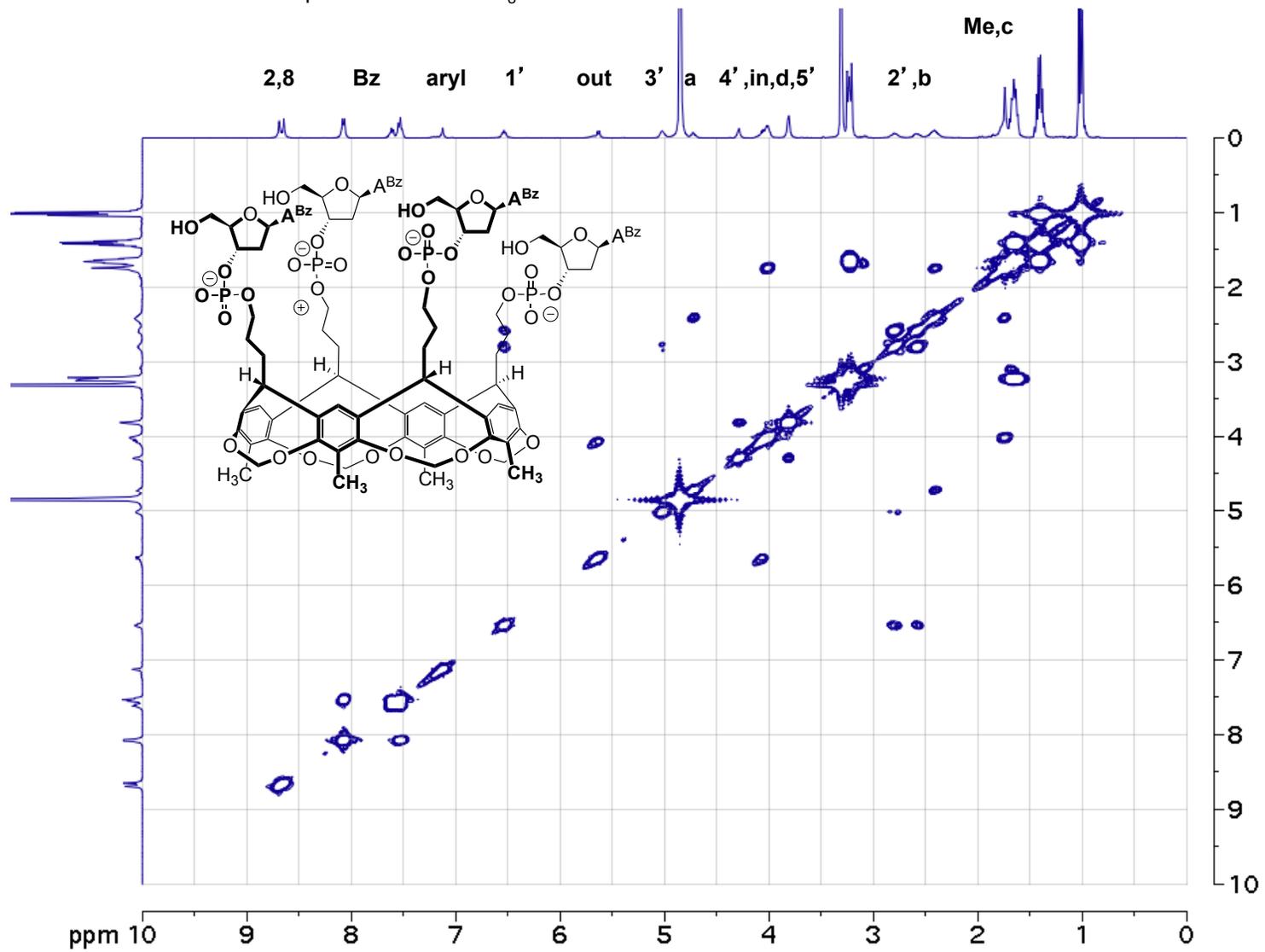
Figure 2.11   ¹H ¹H COSY NMR Spectrum of **9c** in CD₃OD at 400 MHz.
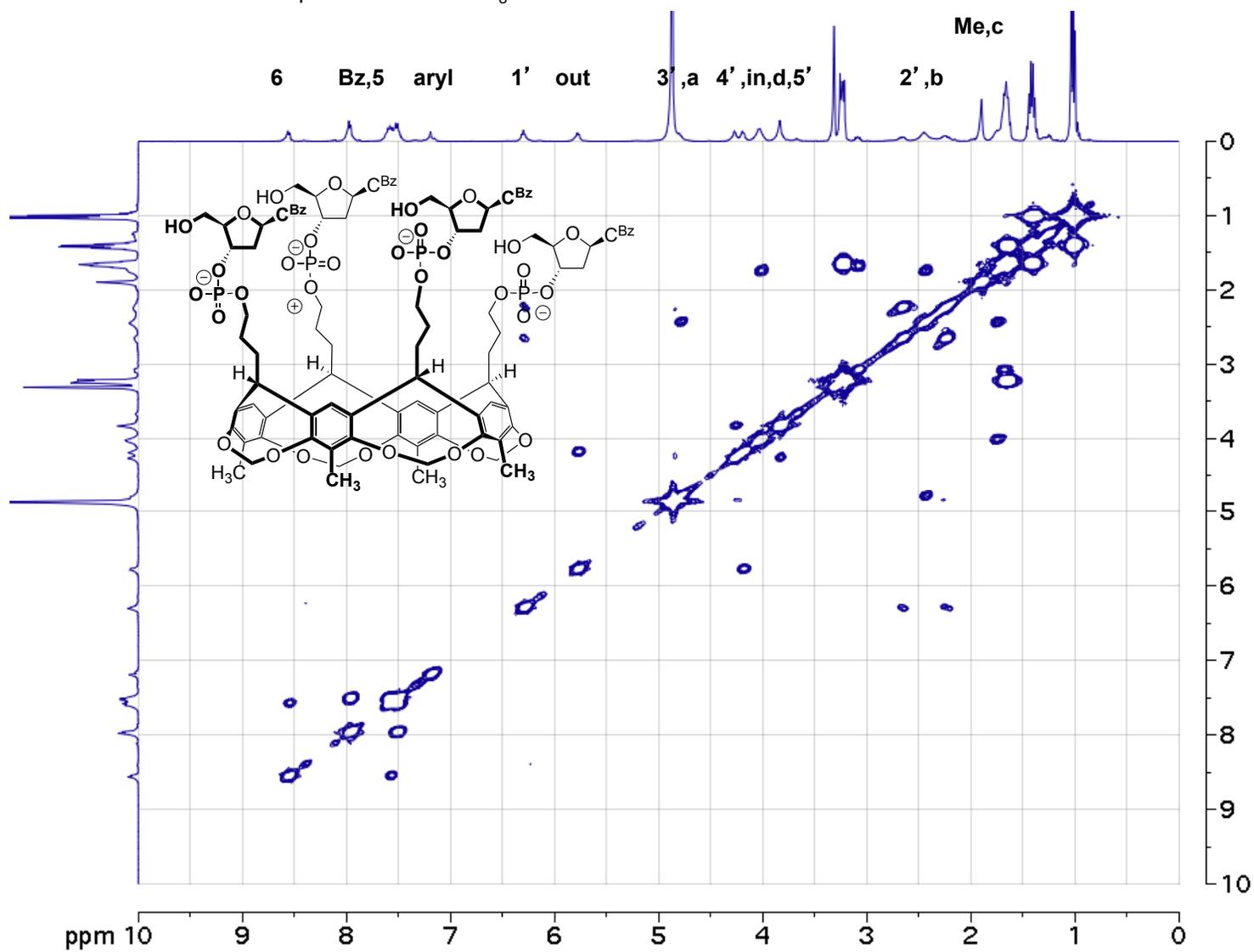
Chapter 3: Synthesis of Water-Soluble Triazole-Linked Cavitand-Guanosine Conjugates

3.1   Synopsis

Two convergent syntheses of a water-soluble cavitand bearing four triazole-linked guanosines were achieved. The critical coupling reaction entailed a copper-catalyzed azide-alkyne cycloaddition between 5'-azido-5'-deoxyguanosine or 5'-azido-5'-deoxy-2',3'-O-isopropylidene-guanosine and a cavitand template functionalized with propargyl ether rim groups and water-solubilizing phosphate pendant groups. This synthesis represented the first known example in the literature of a template bearing four guanosine residues that exhibited water-solubility. The cavitand-guanosine conjugates described here were envisioned to act as template assembled G-quartets and G-quadruplex binding ligand receptors in water. The 2',3'-isopropylidene analogue displayed undesirable broadening in its NMR spectra preventing aquisition of a quality ^{13}C NMR spectrum in $(CD_3)_2SO$ and a quality ^1H NMR spectrum in D_2O. Thus, full characterization of this compound was prevented. The 2',3'-diol analogue demonstrated much improved NMR qualities that allowed its full characterization. A failure to observe any significant signal in the CD spectra in water over a wide range of conditions for these triazole-linked cavitand-guanosine conjugates suggests that these compounds do not offer much potential for G-quartet assembly. Parts of the work described in this chapter have been published already.[123]

3.2 Synthetic Strategy

Pioneering work by Nikan and Sherman demonstrated the utility of a cavitand template to direct the assembly of four covalently linked guanosine residues into a G-quartet in chloroform.[91] Synthesis of this lipophilic template-assembled synthetic G-quartet or TASQ was based on the copper-catalyzed azide-alkyne cycloaddition between 5'-azido-5'-deoxy-2',3'-O-isopropylidene-guanosine nucleoside and a cavitand template functionalized with a propargyl ether rim group to construct a triazole-linked cavitand-guanosine conjugate. The synthetic strategy described here towards a water-soluble analogue of a TASQ was to incorporate water-solubilizing groups at the pendant or 'feet' position of the cavitand while maintaining the triazole-linked guanosine upper rim architecture (Figure 3.1). Functional groups previously shown to impart water-solubility on the resorcin[4]arene based cavitands include alkoxides, carboxylic acids, quaternary ammoniums, and polyethylene glycol chains.[124] Instead, phosphates were chosen as ionizable groups due to their accessibility at the pendant position, their potential to form an octaanionic species, and their existence as the water-solubilizing groups in DNA, RNA, and nucleotides.[115]

Figure 3.1   Strategy for a water-soluble TASQ analogue.

R = triazole-linked rim architecture

R' = water-solubilizing feet
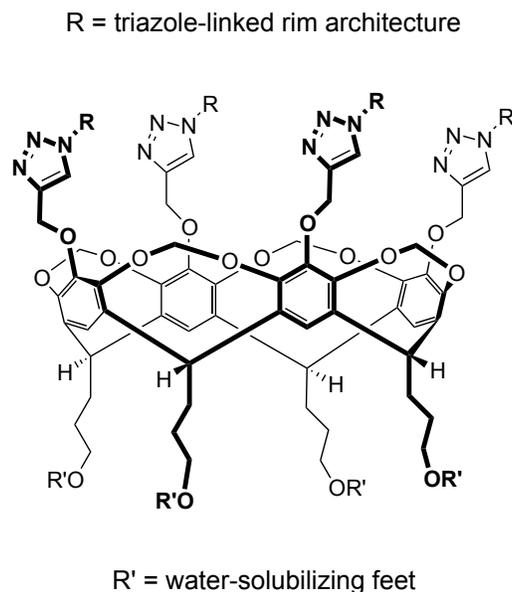
3.3   Synthesis of Triazole-Linked Cavitand-Guanosine Conjugates

A synthetic route already estabilished by Nikan towards propargyl ether rim and phosphate foot functionalized cavitand **16** was employed with modifications.[125] Synthesis of template cavitand **16** required exploitation of recognized orthogonal rim and feet cavitand chemistries (Scheme 3.1).[126,115] Known bromo aryl cavitand **10** was protected at its propyl alcohol feet with *t*-butyldimethylsilylchloride in the presence of imidazole to afford silylated cavitand **11**.[127] To access the upper rim chemistry of **11**, tetrol formation by lithium halogen exchange followed by treatment with trimethyl borate and then

Scheme 3.1 Synthesis of cavitand template.

10

TBS-Cl
imidazole
DMF, 22 h

11
69 %

i) *n*-BuLi, THF, 10 min, -78 °C
ii) B(OMe)$_3$, 3h, -78 °C to rt
iii) NaOH, H$_2$O$_2$, 3h, -78 °C to rt

12
55 %

Br
K$_2$CO$_3$
acetone
reflux, 24h

13
75 %

TBAF
THF, 2h

14
94 %

i) PNEt$_2$(O*t*-Bu)$_2$
   1H-tetrazole
   THF, 1h
ii) H$_2$O$_2$, -78 °C to rt

15
81 %

HO—CF$_3$
DCM, 10 min

16
99 %

≡

76

Scheme 3.1 Synthesis of cavitand template.
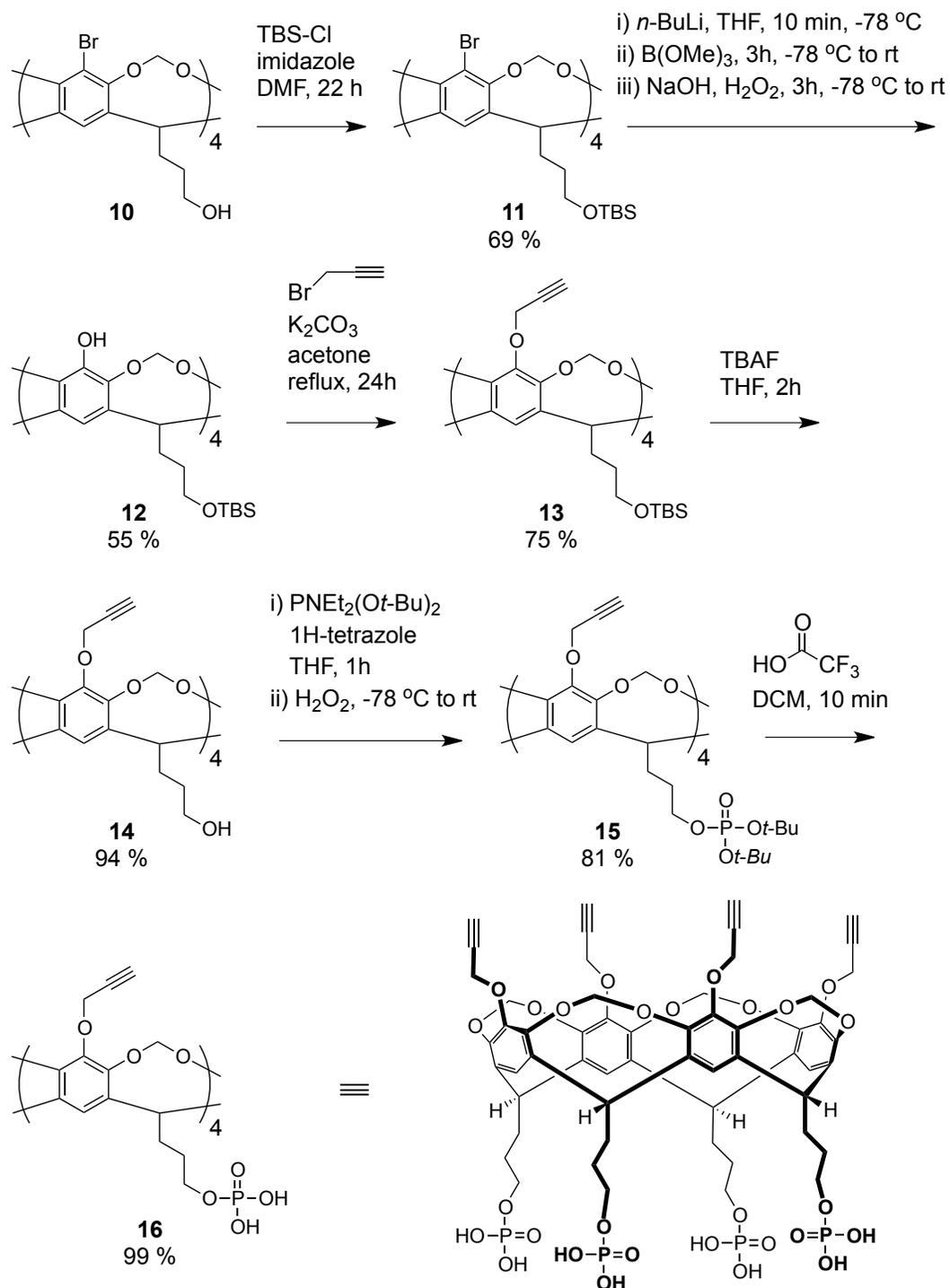
sodium hydroxide / hydrogen peroxide solution was utilized to give phenolic cavitand **12**.[128] Introduction of the alkyne functionality for nucleoside coupling was then performed by refluxing **12** with propargyl bromide and potassium carbonate to give propargyl ether **13**.[129] Removal of the TBS protecting group with fluoride treatment followed by installation of the phosphate group as a protected di-*t*-butyl ester with a 1H-tetrazole activated phosphoramidite reaction yielded cavitand **15**.[115] Final quantitative removal of *t*-butyl by trifluoroacetic acid gave a clean preparation of **16** as a neutral phosphate monoester.

Incorporation of the azido function at the 5' position of guanosine has been reported with tosyl and phosphate activation prior to azide substituion.[130,131] However, these methods are accompanied by intramolecular cyclization with the guanine base at the 5' position and / or poor yields. Dean's improved synthesis based on 5' iodine activation avoids these problems and was utilized to synthesize 5'-azido-5'-deoxyguanosine, **17**.[132] Conversion of nucleoside **17** into its 2',3' isopropylidene variant **18** was accomplished with toluene sulfonic acid and 2,2-dimethoxypropane (Scheme 3.2).[133] Characterization data for **18** from NMR and mass spectrometry were identical to that reported previously.[131]

Scheme 3.2 Nucleosides utilized in the synthesis of triazole linked cavitand-guanosine
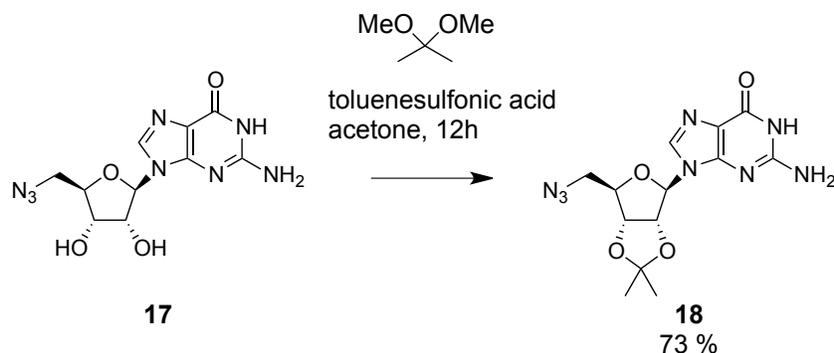
conjugates.

The key step for construction of targets **19a-b** was performed by 'clicking' on nucleosides **17** or **18** via a copper-catalyzed azide-alkyne cycloaddition (Scheme 3.3).[134,135] Final purification of the triazole-linked cavitand-guanosine conjugates was achieved by a classical DNA purification technique with diethylaminoethyl(DEAE)-cellulose anion exchange chromatography and pH 8 triethylammoniumbicarbonate buffer (TEAB) to afford **19a-b** in approximately 30 % yield in triethyammonium salt form.[136] Compounds **19a-b** were found to readily dissolve in the millimolar concentration range in DMSO and water. In opposition to the propensity of guanosine and its derivatives to gel or precipitate in solution, this work demonstrated that a carefully selected template is sufficient to solubilize four guanosine nucleosides on an individual template in the absence of the solubilizing sugar-phosphate backbone present in RNA and DNA oligonucleotides.[137-139] Compounds **19a-b,** in addition to compounds **3a** and **7** described in Chapter 2 represent early examples of water-soluble templates coupled to four guanine containing residues, and were unprecedented in the literature when this work was

undertaken. A peptidic template has been functionalized with triazole-linked guanosine residues and has also demonstrated water-solubility.[140]

3.4 Preliminary Characterization of Triazole-Linked Cavitand-Guanosine Conjugates

Both cavitand-guanosine conjugates described here showed well resolved signals in the ^{1}H NMR spectrum in deuterated DMSO. Full signal assignments could be made under these conditions from two-dimensional NMR correlational spectroscopy ^{1}H ^{1}H COSY experiments (Sections 3.5-6). Diol **19a** demonstrated ideal NMR properties in contrast to isopropylidene **19b** and facilitated the collection of a quality ^{13}C NMR spectrum in $(CD_3)_2SO$ and of a quality ^{1}H NMR spectrum in D_2O. Isopropylidene **19b** gave broadened and uninterpretable NMR spectra for ^{13}C and ^{1}H in $(CD_3)_2SO$ and D_2O, respectively. This broadening was ascribed to aggregation of the compound in solution.[141] Variable temperature NMR experiments were conducted in an effort to denature intermolecular aggregation of **19b** (Figure 3.2) Improved resolution was observed at temperatures of 55 and 85 °C, and diol **19a** is shown in comparison. Circular dichroism (CD) spectroscopy is a highly sensitive technique widely utilized to confirm the existence and conformation of G-quadruplex DNA.[142] Furthermore, lypophilic TASQs were observed to give rise to characteristic CD spectra in chloroform.[92] However, the

Scheme 3.3 Synthesis of triazole-linked cavitand-guanosine conjugates.

16

nucleoside
CuSO$_4$•5H$_2$O
Na ascorbate
DMSO, 20h, 70ºC

nucleoside =

product =

17

R =

19a
31 %

18

R =

19b
34 %
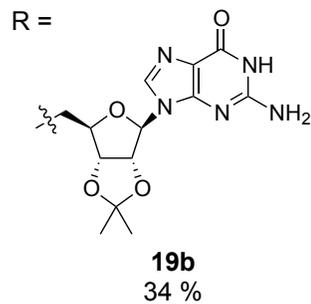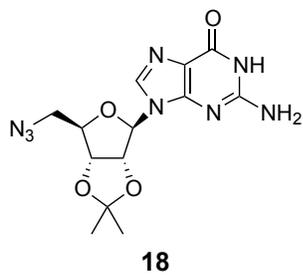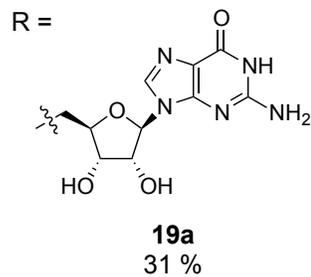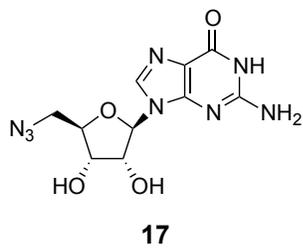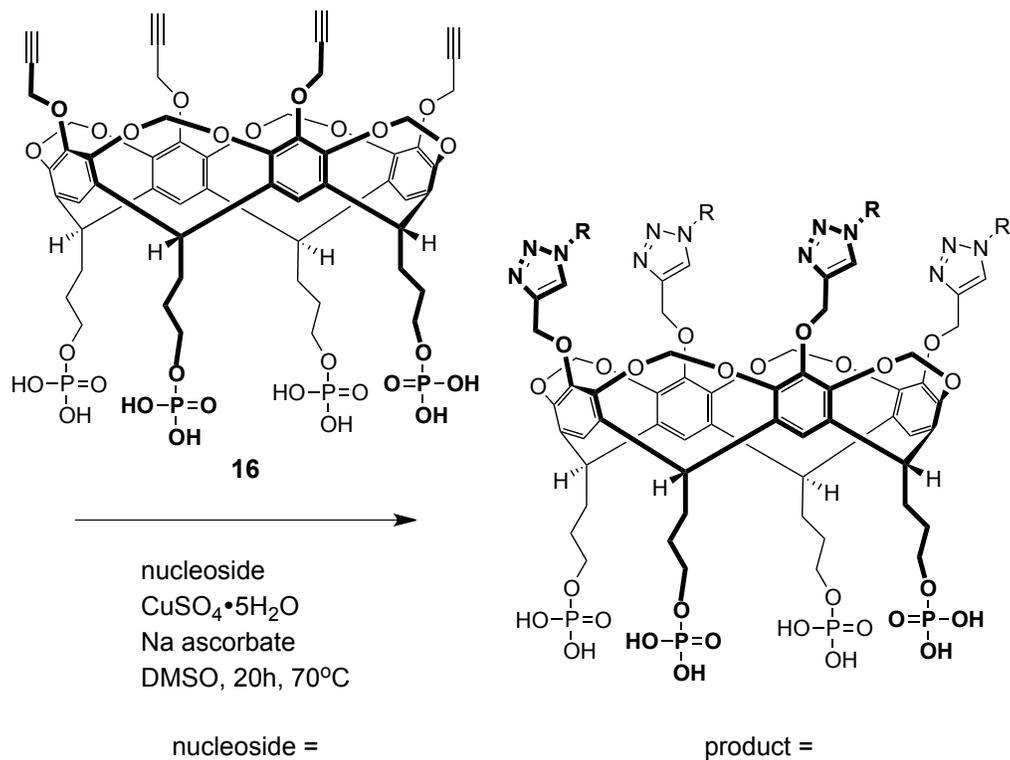
successful acquisition of any significant signal in the CD spectra for compounds **19a-b** over a wide range of conditions, including variations in concentration, cation identity, and pH, in aqueous solution was not possible. This information suggests that the water-soluble triazole-linked cavitand-guanosine conjugates synthesized here do not possess high G-tetrad forming potential.

Figure 3.2   Variable temperature ^{1}H NMR spectra of 2mM solutions in D_2O at 400 MHz for a) diol **19a** and b) isopropylidene **19b**.
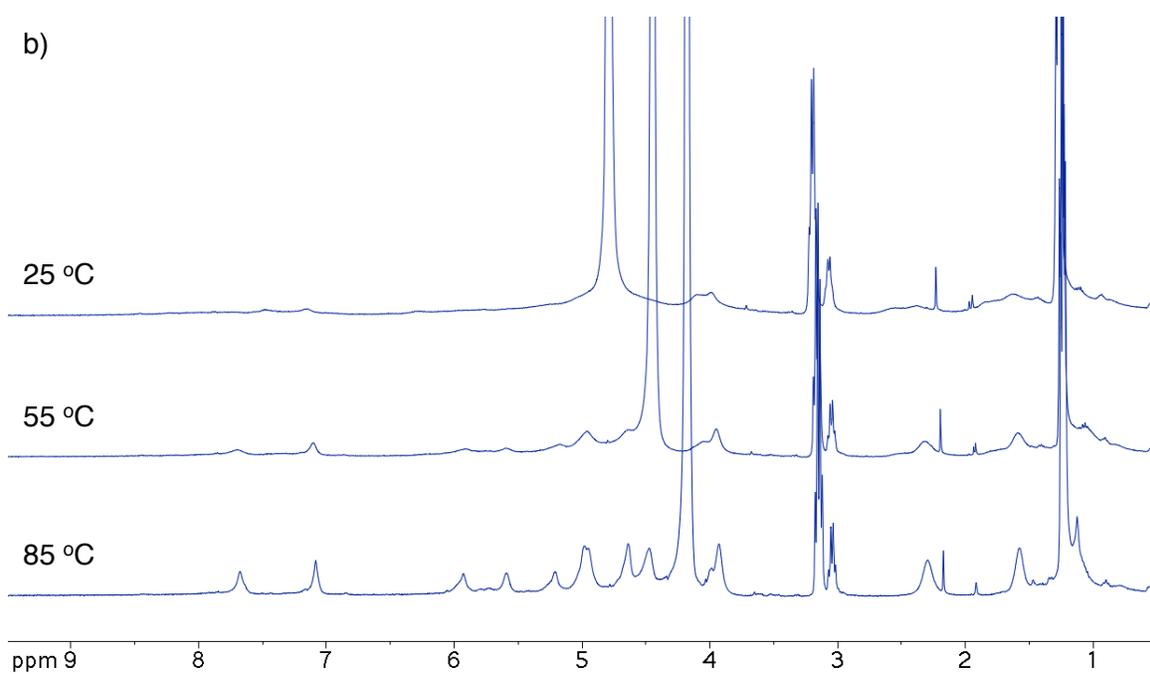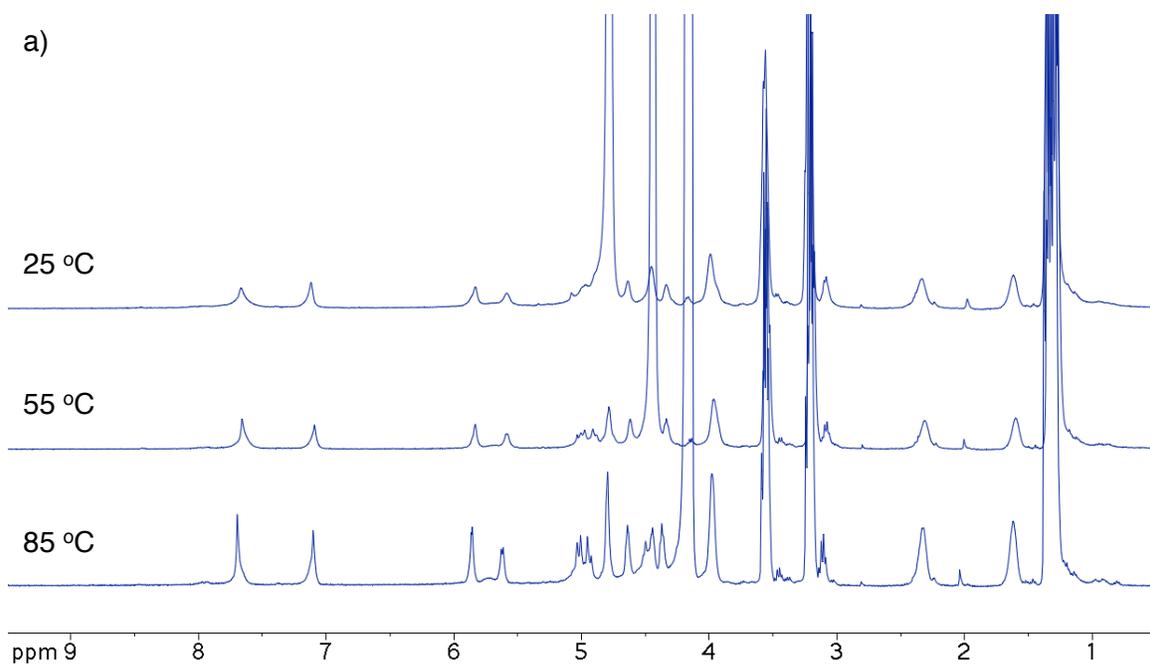
3.5 Experimental Procedures

General Procedures. General procedures outlined in section 2.5 Experimental Procedures were followed with additions given below. Dimethylformamide, toluene, and dimethylsulfoxide were dried over 3 Å molecular sieves. Triethylammonium bicarbonate buffer (TEAB) was prepared by bubbling sublimed CO_2 through a 2 M solution of triethylamine in Milli-Q water at 4 °C until the solution reached pH 8. Synthesis of cavitands **11-16** was based on a synthetic route shown previously.[125] A synthetic route to compound **19b** had been established prior.[143] General numbering conventions used for assignment of NMR spectra are given below.
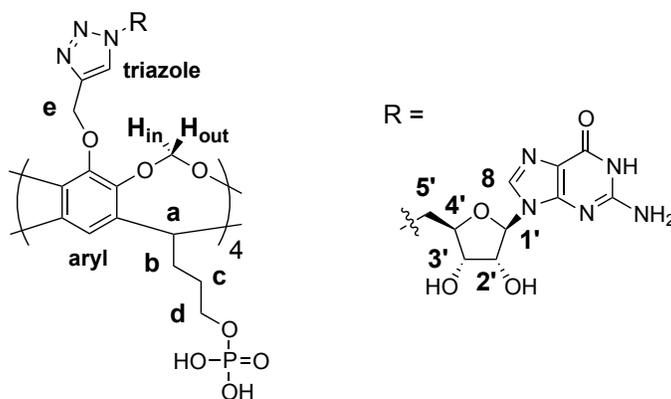
Cavitand 11. To a stirred solution of cavitand **10** (9.88 g, 9.11 mmol, 1 eq) and imidazole (12.42 g, 182.4 mmol, 20 eq) coevaporated two times in 50 mL of dimethylformamide / toluene 1:1 and dissolved in 200 mL of dimethylformamide was added *t*-butyldimethylsilyl chloride (20.63 g, 136.9 mmol, 15 eq) under an argon

atmosphere. The reaction was maintained for 22 hours, 20 mL of water was added, and the reaction was stirred for 30 min further. Solvent was removed by rotary evaporation. The residue was dissolved in 400 mL of chloroform, washed three times with 100 mL of water, and dried with anhydrous $MgSO_4$. After removal of the solvent by rotary evaporation, the residue was dissolved in dichloromethane and embedded on silica. Purification by silica gel column chromatography with ethyl acetate / hexanes 1:50 gave 9.69 g (69 % yield) of **1** as a white solid.

^{1}H NMR (400 MHz, CDCl$_3$) δ = 7.06 (s, 4H, aryl), 5.97 (d, J = 7.5 Hz, 4H, H$_{out}$), 4.89 (t, J = 8.0 Hz, 4H, CH-a), 4.41 (d, J = 7.5 Hz, 4H, H$_{in}$), 3.72 (t, J = 6.5 Hz, 8H, CH$_2$-d), 2.28 (m, 8H, CH$_2$-b), 1.59 (m, 8H, CH$_2$-c), 0.95 (s, 36H, t-butyl), 0.10 (s, 24H, dimethyl); ^{13}C NMR (101 MHz, CDCl$_3$) δ = 152.3, 139.2, 119.1, 113.7, 98.6, 62.5, 37.2, 30.6, 26.2, 26.1, 18.5, -5.1; HRMS (ESI) Calculated for $C_{68}H_{101}O_{12}Si_4Br_4$ (M+H)$^+$: 1539.3083; found: 1539.3123.

Cavitand 12. To a stirred solution of cavitand **11** (3.00 g, 1.95 mmol, 1eq) coevaporated two times in 150 mL of tetrahydrofuran / toluene 1:1, one time with 100 mL of tetrahydrofuran, and dissolved in 450 mL of tetrahydrofuran, was added dropwise n-butyllithium (10 mL, 1.6 M solution in hexanes, 16 mmol, 8.2 eq) under an argon atmosphere at −78 °C. After stirring for 10 minutes, trimethyl borate (2.2 ml, 19 mmol, 9.7 eq) was added, and the solution was allowed to warm to room temperature over 3 hours. The reaction mixture was cooled to −78 °C, and a 90 mL solution of 1.5 M NaOH in 15% H_2O_2 was added. The reaction mixture was allowed to warm to room temperature

over 3 hours and stirring was continued overnight. Excess H_2O_2 was neutralized by the careful addition of $Na_2S_2O_5$ (20.00 g, 105.2 mmol, 54 eq) in portions at 0 °C. The solvent was removed by rotary evaporation, and the residue was suspended in water and suction filtered. Purification by silica gel column chromatography with ethanol / dichloromethane 1:10 then 1:5 gave 1.38 g (55 % yield) of **12** as a white solid.

^{1}H NMR (300 MHz, CDCl$_3$) δ = 6.65 (s, 4H, aryl), 5.97 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.74 (t, J = 8.0 Hz, 4H, CH-a), 4.45 (d, J = 7.0 Hz, 4H, H$_{in}$), 3.71 (t, J = 6.5 Hz, 8H, CH$_2$-d), 2.24 (m, 8H, CH$_2$-b), 1.59 (m, 8H, CH$_2$-c), 0.95 (s, 36H, t-butyl), 0.09 (s, 24H, dimethyl), 4 phenolic protons not observed due to exchange; ^{13}C NMR (101 MHz, CDCl$_3$) δ = 142.2, 141.0, 138.5, 110.2, 99.9, 62.8, 36.4, 30.8, 26.2, 25.9, 18.6, -5.1; HRMS (ESI) Calculated for $C_{68}H_{105}O_{16}Si_4$ (M+H)$^+$: 1289.6480; found: 1289.6467.

Cavitand 13. To a stirred suspension of cavitand **12** (1.10 g, 0.853 mmol, 1 eq) and anhydrous potassium carbonate (2.40 g, 17.4 mmol, 20 eq) in 60 mL of acetone was added propargyl bromide (1.5 mL, 80 % w/w propargyl bromide in toluene, 14 mmol, 16 eq) under an argon atmosphere. The reaction mixture was refluxed for 24 hours. The solvent was removed by rotary evaporation. The residue was suspended in 100 mL of dichloromethane and stirred for 30 minutes at 35 °C. The resultant precipitate was removed by gravity filtration and the mother liquor was concentrated. Purification by silica gel column chromatography with methanol / dichloromethane 1:100 gave 0.92 g (75 % yield) of **13** as an off-white solid.

^{1}H NMR (300 MHz, CDCl$_3$) δ = 6.82 (s, 4H, aryl), 5.90 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.75 (t, J = 8.0 Hz, 4H, CH-a), 4.61 (d, J = 2.5 Hz, 8H, CH$_2$-e), 4.40 (d, J = 7.0 Hz, 4H, H$_{in}$), 3.71 (t, J = 6.5 Hz, 8H, CH$_2$-d), 2.48 (t, J = 2.5 Hz, 4H, alkyne), 2.23 (m, 8H, CH$_2$-b), 1.60 (m, 8H, CH$_2$-c), 0.95 (s, 36H, t-butyl), 0.09 (s, 24H, dimethyl); ^{13}C NMR (101 MHz, CDCl$_3$) δ = 148.4, 143.7, 138.8, 114.8, 99.7, 79.6, 75.0, 63.0, 61.0, 36.4, 30.9, 26.2, 26.1, 18.6, -5.0; HRMS (ESI) Calculated for C$_{80}$H$_{113}$O$_{16}$Si$_4$ (M+H)$^+$: 1441.7106; found 1441.7133.

Cavitand 14. To a stirred solution of cavitand **13** (900 mg, 0.624 mmol, 1 eq) in 40 mL of tetrahydrofuran was added tetrabutylammonium fluoride (7.5 mL, 1.0 M solution in tetrahydrofuran, 7.5 mmol, 12 eq) at 0 °C. The reaction was warmed to room temperature and maintained for 2 hours. The solvent was removed by rotary evaporation, and the residue was embedded on silica gel. Purification by silica gel column chromatography with ethanol / ethyl acetate 1:10 gave 578 mg (94 % yield) of **14** as a white solid.

^{1}H NMR (400 MHz, CDCl$_3$) δ = 6.98 (s, 4H, aryl), 5.91 (d, J = 7.0 Hz, 4H, H$_{out}$), 4.76 (t, J = 8.0 Hz, 4H, CH-a), 4.62 (d, J = 2.5 Hz, 8H, CH$_2$-e), 4.42 (d, J = 7.0 Hz, 4H, H$_{in}$), 3.79 (t, J = 5.5 Hz, 8H, CH$_2$-d), 2.49 (t, J = 2.5 Hz, 4H, alkyne), 2.39 (m, 8H, CH$_2$-b), 1.63 (br, 8H, CH$_2$-c), 1.26 (br, 4H, OH); ^{13}C NMR (101 MHz, (CD$_3$)$_2$SO) δ = 147.4, 142.8, 138.7, 116.6, 99.2, 79.6, 78.2, 60.4, 60.4, 36.7, 30.9, 25.4; HRMS (ESI) Calculated for C$_{56}$H$_{57}$O$_{16}$ (M+H)$^+$: 985.3647; found: 985.3677.

Cavitand 15. To a rapidly stirred suspension of cavitand **14** (125 mg, 0.127 mmol, 1 eq) and 1H-tetrazole (284 mg, 4.05 mmol, 32 eq) coevaporated two times in 2 mL of

tetrahydrofuran and dissolved in 10 mL of tetrahydrofuran was added dropwise di-*t*-butyl *N,N*-diethylphosphoramidite (0.58 mL, 2.1 mmol, 16 eq) under an argon atmosphere. The reaction was maintained for 1 hour. The reaction was cooled to -78 °C, and 0.4 mL of 30 % H_2O_2 was added. The reaction was allowed to warm to room temperature, and the solvent was removed by rotary evaporation. The residue was dissolved and extracted three times with 30 mL of dichloromethane from 50 mL of aqueous 5% $Na_2S_2O_5$. The combined organic phase was washed with 50 mL of saturated $NaHCO_3$ and 50 mL of brine, dried over anhydrous $MgSO_4$, and the solvent was removed. Purification by silica gel column chromatography with methanol / dichloromethane / triethylamine 1:50:0.5 gave 180 mg (81 % yield) of **15** as a glassy solid.

^{1}H NMR (400 MHz, CDCl$_3$) δ = 6.82 (s, 4H, aryl), 5.89 (d, *J* = 7.5 Hz, 4H, H$_{out}$), 4.79 (t, *J* = 8.5 Hz, 4H, CH-a), 4.60 (d, *J* = 2.5 Hz, 8H, CH$_2$-e), 4.39 (d, *J* = 7.5 Hz, 4H, H$_{in}$), 4.04 (m, 8H, CH$_2$-d), 2.48 (t, *J* = 2.5 Hz, 4H, alkyne), 2.33 (m, 8H, CH$_2$-b), 1.73 (m, 8H, CH$_2$-c), 1.49 (s, 72H, *t*-butyl); ^{13}C NMR (101 MHz, CDCl$_3$) δ = 148.5, 143.7, 138.6, 114.8, 99.6, 82.2, 79.5, 75.1, 66.3, 60.9, 36.3, 30.0, 29.9, 25.7; HRMS (ESI) Calculated for $C_{88}H_{124}O_{28}P_4Na$ (M+Na)$^+$: 1775.7127; found: 1775.7097.

Cavitand 16. To a stirred solution of cavitand **15** (139 mg, 79.3 umol, 1 eq) coevaporated two times in 2 mL of dichloromethane and dissolved in 5 mL of dichloromethane was added trifluoracetic acid (1.2 mL, 16 mmol, 200 eq) under an argon atmosphere. The reaction was maintained for 10 minutes. Removal of solvent by rotary evaporation and

coevaporation three times with 5 mL of methanol gave 102 mg (99 % yield) of **16** as a white solid.

^{1}H NMR (300 MHz, CD$_3$OD) δ = 7.10 (s, 4H, aryl), 5.86 (d, J = 7.5 Hz, 4H, H$_{out}$), 4.75 (t, J = 8.0 Hz, 4H, CH-a), 4.61 (d, J = 2.5 Hz, 8H, CH$_2$-e), 4.38 (d, J = 7.5 Hz, 4H, H$_{in}$), 4.11 (br, 8H, CH$_2$-d), 2.89 (t, J = 2.5 Hz, 4H, alkyne), 2.42 (m, 8H, CH$_2$-b), 1.71 (br, 8H, CH$_2$-c), 8 phosphoric protons not observed due to exchange; ^{13}C NMR (101 MHz, CD$_3$OD) δ = 149.9, 144.8, 139.8, 116.6, 101.3, 80.1, 76.7, 67.6, 61.7, 38.2, 30.2, 26.9; HRMS (ESI) Calculated for C$_{56}$H$_{59}$O$_{28}$P$_4$ (M-H)$^-$: 1303.2143; found: 1303.2102.

Nucleoside 18. To a stirred suspension of nucleoside **17** (1.00 g, 3.23 mmol, 1 eq) and *p*-toluenesulfonic acid monohydrate (0.42 g, 3.9 mmol, 1.2 eq) in 25 mL of acetone was added 2,2-dimethoxypropane (4.0 mL, 33 mmol, 10 eq). The reaction was maintained for 12 hours. Concentrated aqueous ammonium hydroxide solution was added dropwise until the reaction mixture turned basic, and the resultant white precipitate was suction filtered. The precipitate was partitioned between 50 mL of chloroform and 50 mL of saturated aqueous NaHCO$_3$ and extracted two times. The combined organic phase was washed with 50 mL of brine, dried over anhydrous MgSO$_4$, and embedded on silica gel. Purification by silica gel column chromatography with ethanol / chloroform / triethylamine 1:4:0.05 gave 0.82 g (73 % yield) of **18** as an off-white solid.

General click synthetic procedure for triazole linked cavitand-nucleoside conjugates 19a-b. To a stirred solution of cavitand **16** (20 mg, 15 umol, 1 eq) and nucleoside **17** or

18 (66 umol, 4.4 eq) in 10 mL of argon purged dimethylsulfoxide was added a solution of copper(II) sulfate pentahydrate (0.65 mL, 0.008 M in argon purged Mili-Q water, 0.52 umol, 0.35 eq) and a solution of sodium ascorbate (1.3 mL, 0.04 M in argon purged Mili-Q water, 5.2 umol, 3.5 eq) under an argon atmosphere. The reaction was heated to 70 °C and maintained for 20 hours. Solvent was removed by rotary evaporation, and the residue was suspended in water, sonicated, and suction filtered. The resulting precipitate was dissolved in a 5 mL mixture of triethylamine / water 4:1, vortexed, and lyophilized. The crude product was dissolved in 0.01 M TEAB buffer and purified by ion-exchange chromatography on DEAE-cellulose (30 cm x 1.5 cm column, detection at 254 nm) with gradient elution (3 mL min^{-1}, 0.5 M to 1 M TEAB buffer, 50 mins) followed by isocratic elution (1 M TEAB buffer). Reduction of the solvent followed by lyophilization afforded **19a-b** in triethylammonium form as white powders.

19a. 16 mg (31 % yield); triethylamine was removed for collection of NMR spectra in $(CD_3)_2SO$ by repeated coevaporation with $(CD_3)_2SO$: ^{1}H NMR (400MHz, $(CD_3)_2SO$) δ = 8.08 (s, 4H, triazole), 7.84 (s, 4H, CH-8'), 7.37 (s, 4H, aryl), 6.53 (br, 8H, NH$_2$), 5.86 (d, $J = 7.5$ Hz, 4H, H$_{out}$), 5.70 (d, $J = 6.5$ Hz, 4H, CH-1'), 4.94 (s, 8H, CH$_2$-e), 4.74 (m, 8H, CH$_2$-5'), 4.55 (m, 8H, CH-2', CH-a), 4.24 (m, 8H, CH-4', H$_{in}$), 4.18 (m, 4H, CH-3'), 3.92 (br, 8H, CH$_2$-d), 2.38 (br, 8H, CH$_2$-b), 1.53 (br, 8H, CH$_2$-c), 20 protons not observed due to water suppression and / or exchange; ^{1}H NMR (400MHz, D$_2$O) δ = 7.66 (br, 8H, triazole, CH-8), 7.11 (s, 4H, aryl), 5.81 (br, 4H, CH-1'), 5.53 (br, 4H, H$_{out}$), 5.3-4.7 (overlap, 16H, CH$_2$-e, CH$_2$-5'), 4.61 (br, 4H, CH-2'), 4.43 (br, 8H, CH-a, CH-4'), 4.30 (br, 4H, CH-3'), 3.95 (br, 12H, CH$_2$-d, H$_{in}$), 2.32, (br, 8H, CH$_2$-b), 1.60, (br, 8H, CH$_2$-c),

28 protons not observed due to exchange; ^{13}C NMR (101 MHz, (CD$_3$)$_2$SO) δ = 156.8, 153.6, 151.2, 147.5, 143.8, 143.3, 138.7, 136.2, 125.2, 117.0, 116.3, 99.3, 87.0, 82.5, 72.4, 71.1, 66.5, 64.8, 51.6, 37.5, 29.3, 25.6;); ^{31}P NMR (162 MHz, (CD$_3$)$_2$SO) δ = 1.1; MS (MALDI-TOF) calculated for C$_{96}$H$_{107}$N$_{32}$O$_{44}$P$_4$ (M-H)$^-$: 2535.6; found: 2537.8; HRMS (ESI) calculated for C$_{96}$H$_{105}$N$_{32}$O$_{44}$P$_4$ (M-3H)$^{-3}$: 844.5304; found: 844.5305.

19b. 18 mg (34 % yield); ^{1}H NMR (400MHz, (CD$_3$)$_2$SO) δ = 8.03 (s, 4H, triazole), 7.83 (s, 4H, CH-8), 7.35 (s, 4H, aryl), 6.74 (br, 8H, NH$_2$), 6.06 (br, 4H, CH-1'), 5.81 (br, 4H, H$_{out}$), 5.28 (br, 8H, CH-2', CH-3'), 4.98 (m, 8H, CH$_2$-e), 4.67 (br, 8H, CH$_2$-5'), 4.49 (br, 8H, CH-a, CH-4'), 4.22 (br, 4H, H$_{in}$), 3.86 (br, 8H, CH$_2$-d), 2.33 (overlap, 8H, CH$_2$-b), 1.47 (s, 12H, methyl), 1.47 (overlap, 8H, CH$_2$-c), 1.30 (s, 12H, methyl); MS (MALDI-TOF) calculated for C$_{108}$H$_{125}$N$_{32}$O$_{44}$P$_4$ (M+H)$^+$: 2697.7; found: 2699.5.

3.6   Supplementary ^{1}H ^{1}H COSY NMR spectra

Supplementary ^{1}H ^{1}H COSY NMR spectra begins on the following page.

Figure 3.3   ¹H ¹H COSY NMR spectrum of **19a** in $(CD_3)_2SO$ at 400 MHz.
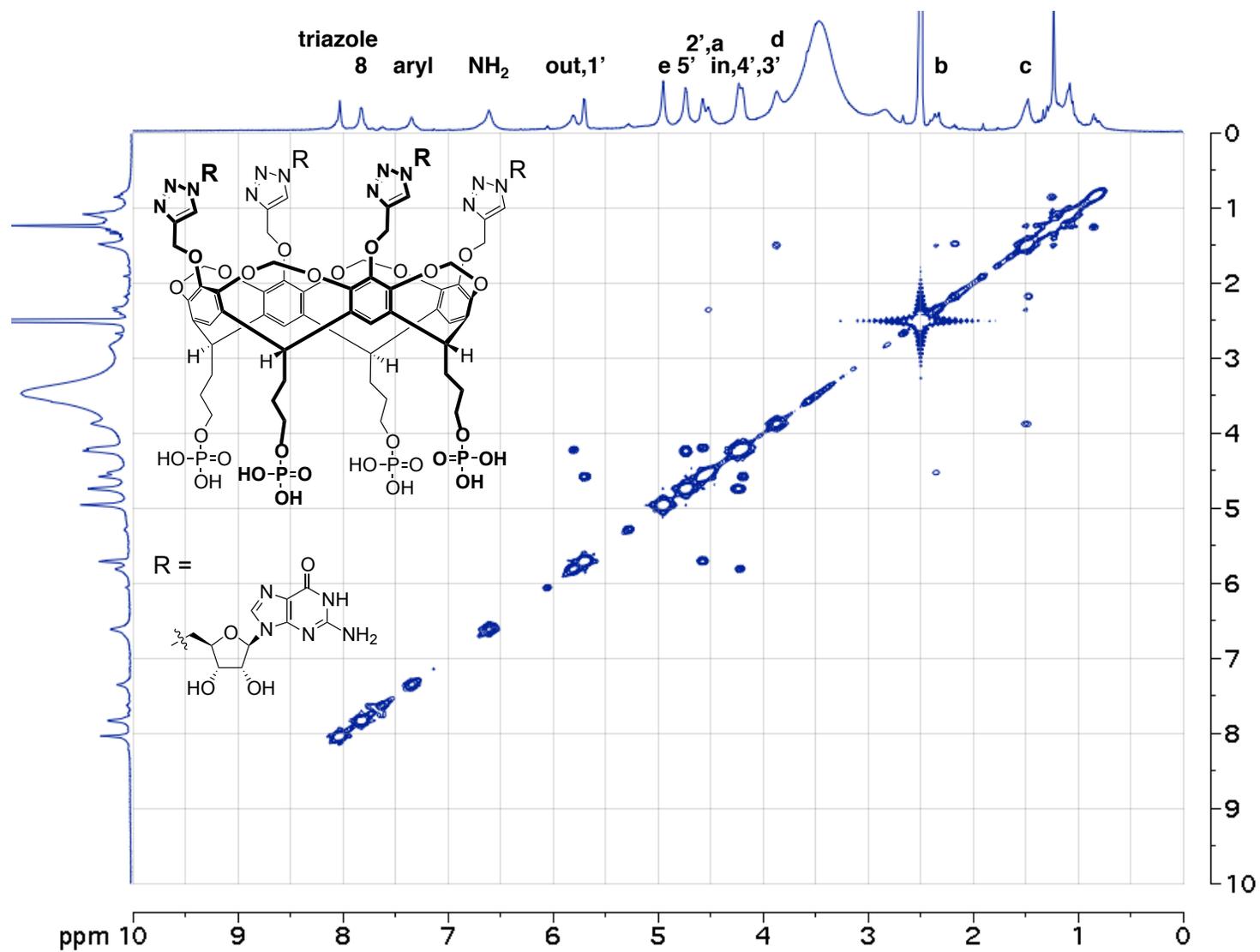
Figure 3.4   ¹H ¹H COSY NMR spectrum of **19a** in D_2O at 400 MHz.
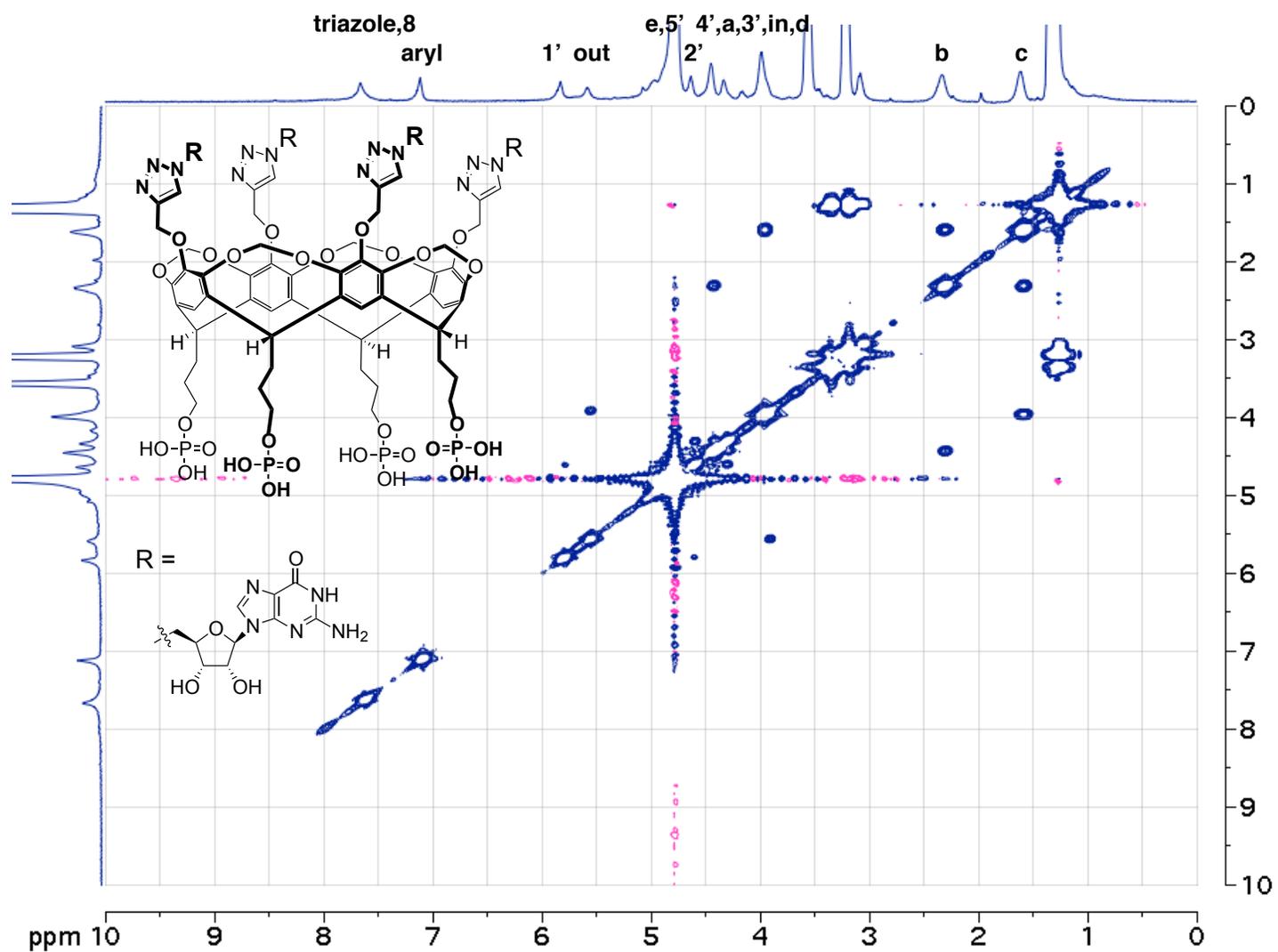
Figure 3.5 ^{1}H ^{1}H COSY NMR spectrum of **19b** in (CD$_3$)$_2$SO at 400 MHz.
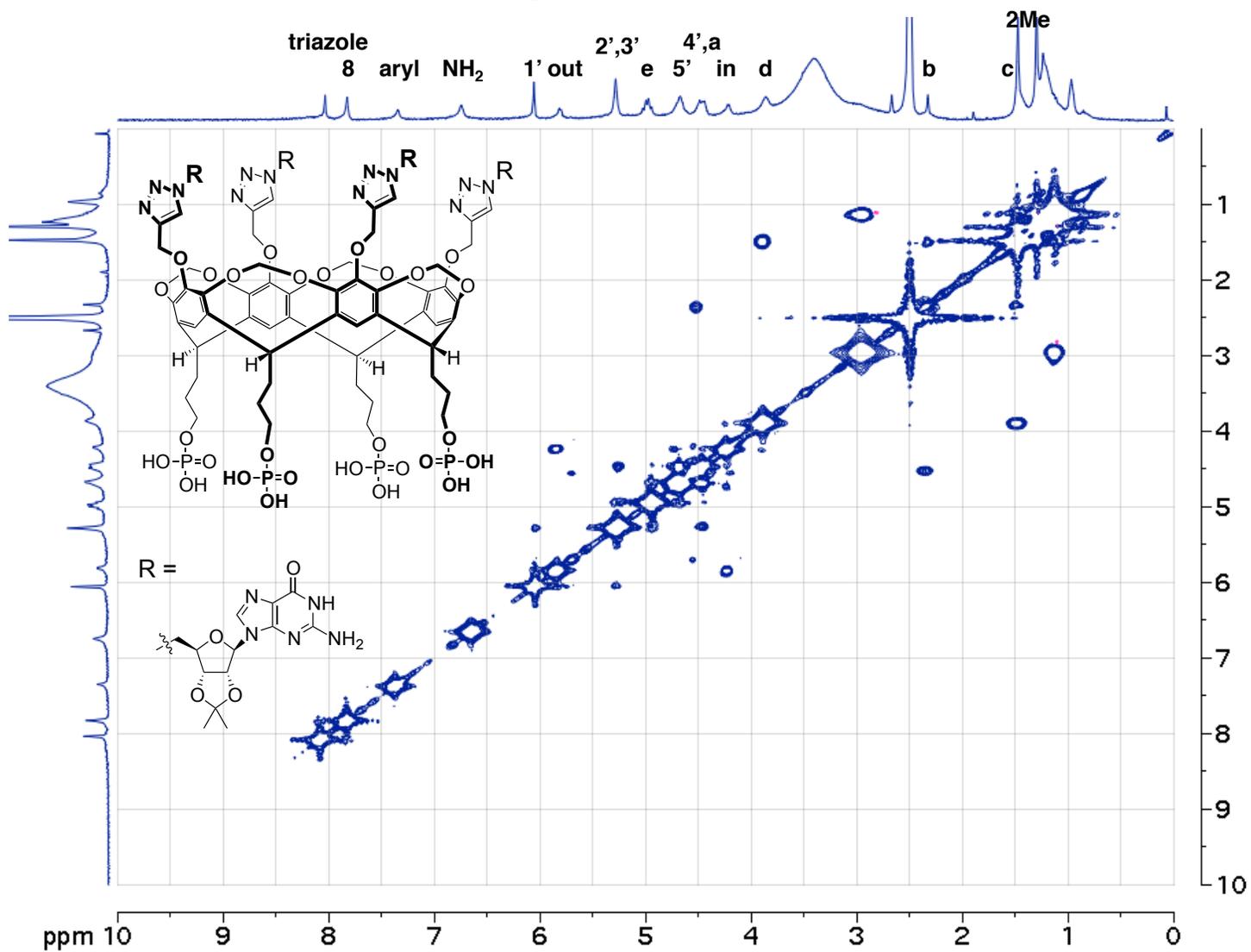
Chapter 4: Discovery of Cation-Free Unimolecular Template-Assembled Guanine and Thymine Quartets in Methanol

4.1   Synopsis

Phosphate-linked cavitand-nucleotide conjugates functionalized at the feet position were evaluated for template-assembly of the corresponding nucleobases into hydrogen bonded quartet structures in methanol solvent. The conjugates were initially scanned by CD spectroscopy. Spectra of the conjugates in methanol were compared to spectra in a potentially denaturing aqueous solution and to spectra of base-protected analogue compounds. These data were used to argue that potential conjugate template-assembly was observed with the unprotected conjugate in methanol. Adenine and cytosine based cavitand-nucleotide conjugates displayed weak or no CD differences under the above conditions, and their characterization was not rigorous. However, guanine and thymine conjugates exhibited CD in methanol that was suppressed significantly in water or in the case of the guanine conjugate by blocking the base. Additional NMR characterization by proton 1D, NOESY, and ROESY experiments demonstrated that both guanine and thymine conjugates possessed downfield imino resonances with increased kinetic stability typically found for hydrogen bonded base pairs and contained *syn* aligned nucleobases. An additional ROE correlation between the imino and 5' methyl signal was found for the thymine conjugate in pure methanol, and

this established increasing support for a rare template-assembled thymine tetrad. Finally, DOSY NMR confirmed that the guanine and thymine conjugates exist as unimolecular species in solution. The methodologies employed in the synthesis of the conjugates here were shown not to be accompanied by any significant amount of sodium contamination by ICP-MS. All of this evidence together builds a strong case for the discovery of cation-free unimolecular template-assembled guanine and thymine quartets in methanol.

4.2 Circular Dichroism of Phosphate-Linked Cavitand-Nucleotide Conjugates at Feet Position

Circular dichroism (CD) is a type of chiroptical spectroscopy whereby a non-racemic sample of a chiral compound exhibits differential absorption between left and right circularly polarized light.[144] In addition to its use for the assignment of the absolute configurations in organic molecules, this highly sensitive technique has found widespread application in the study of nucleic acid and protein secondary structure.[145,146] Analysis of G-quadruplex folding topology, which includes the elements of strand sense and stoichiometry, glycosidic bond orientation, and loop conformation, by CD is no exception.[142,147-149] The interpretation of CD spectra of G-quadruplex DNA to date has been almost exclusively empirical with two G-quadruplex CD signatures being dominant in the literature. Parallel G-quadruplexes with all *anti* glycosidic bond conformations are characterized by respective positive and negative bands near to 260 and 240 nm. In

comparison, antiparallel G-quadruplexes with alternating *anti / syn* glycosidic bond conformations are characterized by positive and negative bands at approximately 290 and 260 nm, respectively.

The major CD bands characterizing G-quadruplex DNA have been explained recently to arise from exciton coupling between nearest stacked guanine bases in vertically adjacent planar G-tetrads (Figure 4.1).[142] In this model, CD generating chiral

Figure 4.1   An exciton coupling interaction between electric transition dipoles (black) of two
adjacently stacked guanine bases in G-quadruplex DNA.
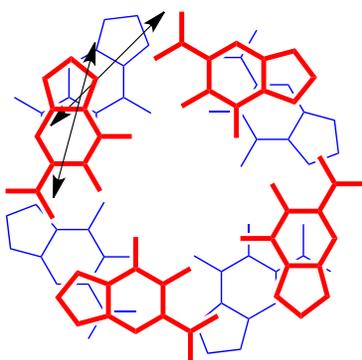
exciton couplings occur from the interaction between electric transition dipoles of intrinsically non-chiral guanine chromophores that are orientated close in space and at angles defined by the helical twist of layered G-tetrads. When one chromophore is located close in space and chirally oriented to a second chromophore, the resultant electric transition dipole interaction can give a CD enhancement significantly beyond that originating from the interaction between a chromophore and a non-chromophoric asymmetric carbon skeleton.[144] The latter case is represented by the CD of nucleosides,

which are chiral molecules made up from a planar nucleobase chromophore and a furanose ring with high stereocenter density.[150] Early studies involving CD spectroscopy of G-quadruplex DNA demonstrated that an induction and enhancement of the signature CD signal accompanied increasing potassium stabilization of the G-quadruplex.[151,152] Loss of CD signal for potassium stabilized G-quadruplex DNA was correlated with thermal denaturation of the folded quadruplex secondary structure.

In consideration of the above discussion, a unimolecular guanine or other nucleobase quartet, such as that formed from a template-assembled system or that arising from free nucleoside association, should give rise to an induced CD through chiral exciton couplings between neighboring bases within the quartet plane (Figure 4.2).

Figure 4.2   Some potential exciton couplings in the two helical configurations of a single

nucleobase quartet.
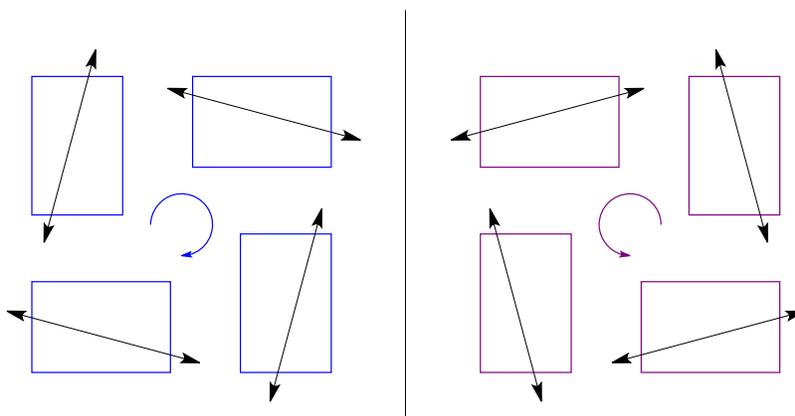
A clockwise or a counterclockwise helical arrangement for a single nucleobase quartet is possible. Helically arranged bases of a single quartet would not however be expected to

97

be empirically interpretable relative to existing G-quadruplex CD spectra for the reason that base-stacked or strand-connected nucleosides are absent in these systems i.e., information about *syn* or *anti* dihedral angles cannot be inferred from these spectra.[142] Some known and potential base quartets of the DNA bases are given (Figure 4.3).

Figure 4.3   Nucleobase quartets for a) guanine (common), b) adenine (rare), c) thymine (rare), and d) cytosine (not known).
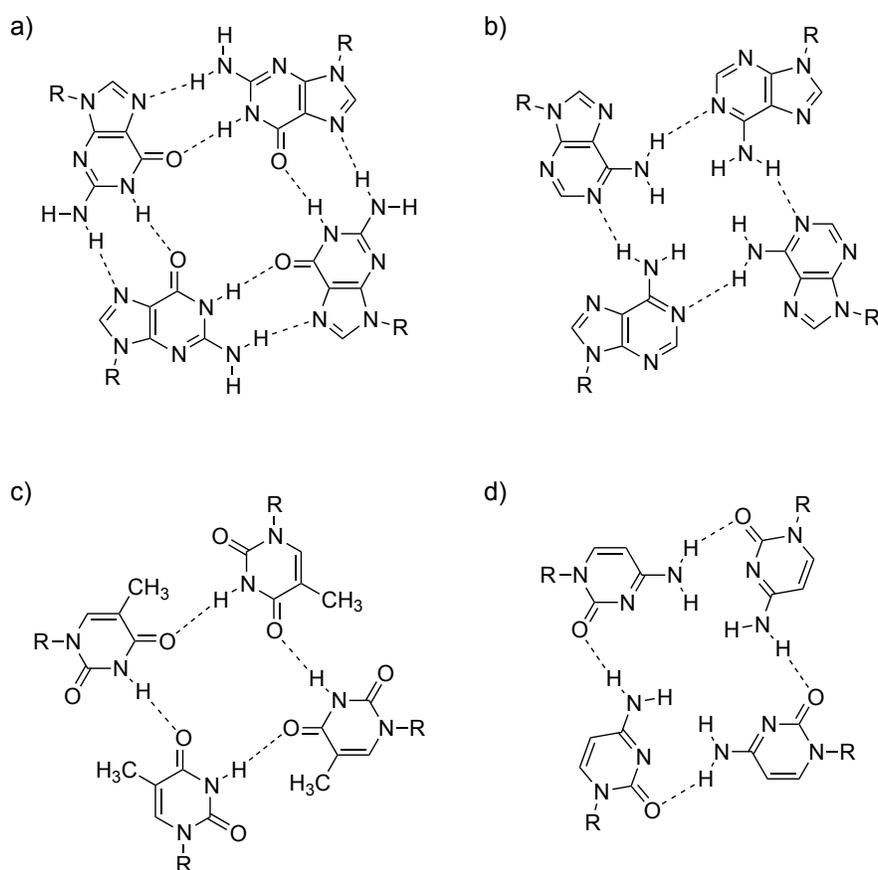
The requirement of the nucleobase to be part of a nucleoside or nucleotide should be emphasized here, since a quartet of the nucleobase alone is not sufficient to remove

98

mirror plane symmetry. Chloroform soluble guanine and uracil template-assembled synthetic quartets are unimolecular in the absence of cations and display characteristic CD signatures.[92,103] Thus, CD can be considered a useful diagnostic tool for the investigation of whether template-assembly is taking place. This conclusion can be aided by comparison to analogue compounds with restricted capacity for template-assembly or by comparison under denaturing conditions.

Accordingly, the propensity of methanol and water-soluble phosphate-linked cavitand-nucleotide conjugates **3a-d** and **9a-c** to template-assemble nucleotide quartets was investigated (Figure 4.4). Blocking the base exocyclic amino function, a potential hydrogen bond donating group, with benzoyl or isobutyryl protection was expected to accompany a loss of self-assembly for compounds **9a-c**, and these blocked analogues were expected to show depressed CD spectra. As well, suppression of CD for unblocked cavitand-nucleotide conjugates **3a-d** was supposed in a denaturing solvent, if one could be found. Thus, an enhancement in the CD spectrum of a cavitand-nucleotide conjugate relative to its blocked analogue or relative to a second solvent suggests that a self-assembly process, like that for a template-assembled nucleotide quartet, is active.

Circular dichroism spectra of cavitand-nucleotide conjugates dissolved in methanol or aqueous buffer were recorded along with spectra for the blocked analogue conjugates. Benzoyl protected compounds **9b-c** did not demonstrate sufficient water solubility for comparison. In the case of adenine derived cavitand-nucleotide conjugates **3b** and **9b**, poor signal to noise ratios were apparent. Significant magnitude differences were absent in the CD spectra for the unblocked conjugate **3b** with respect to benzoyl protected conjugate **9b** or between different solvents (Figure 4.5). A weak negative CD

that resides at approximately 280 nm for **9b** and 265 nm and 255 nm for **3b** in methanol

and water respectively may be present in these spectra.

Figure 4.4   Phosphate-linked cavitand-nucleotide conjugates at feet position.

3a, R = G

9a, R = G^Ibu

3b, R = A

9b, R = A^Bz

3c, R = C

9c, R = C^Bz

3d, R = T
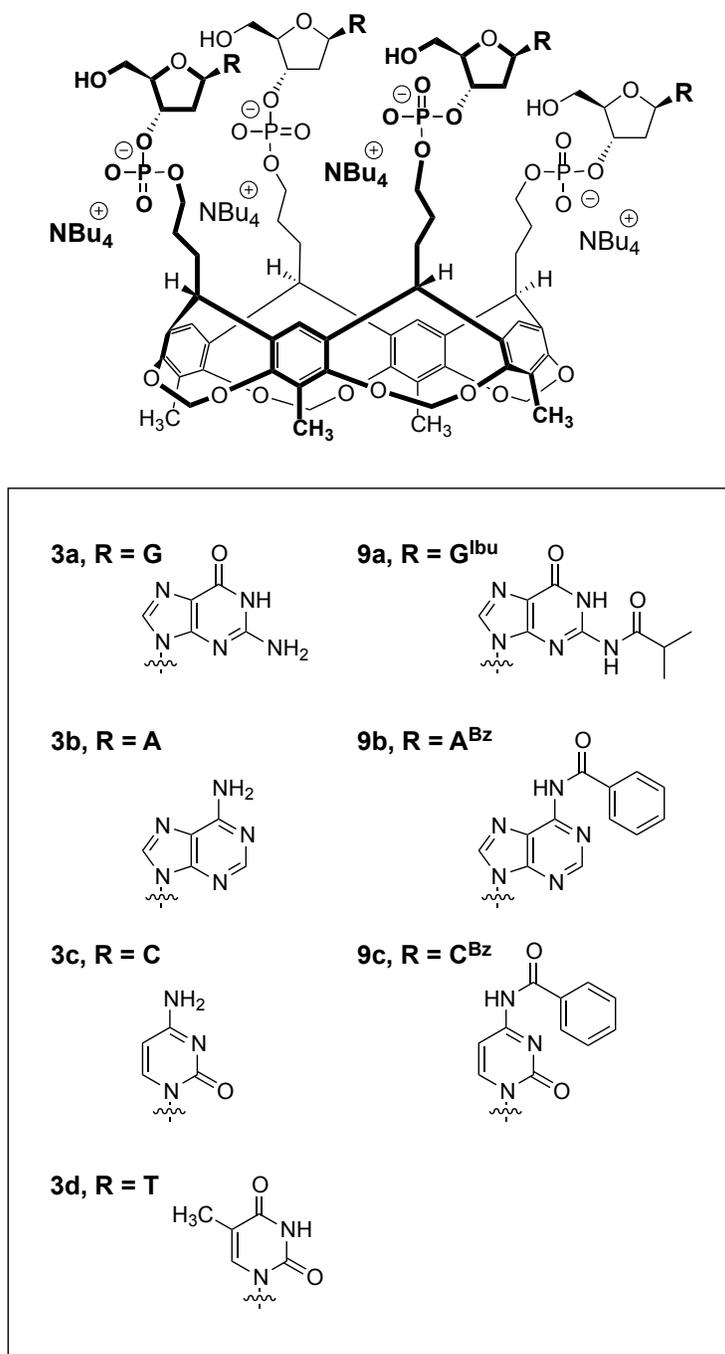
Figure 4.5   CD spectra of adenine conjugate **3b** in methanol (red) and in 10 mM Tris-borate pH 8 buffer (blue) and of blocked adenine conjugate **9b** in methanol (green).
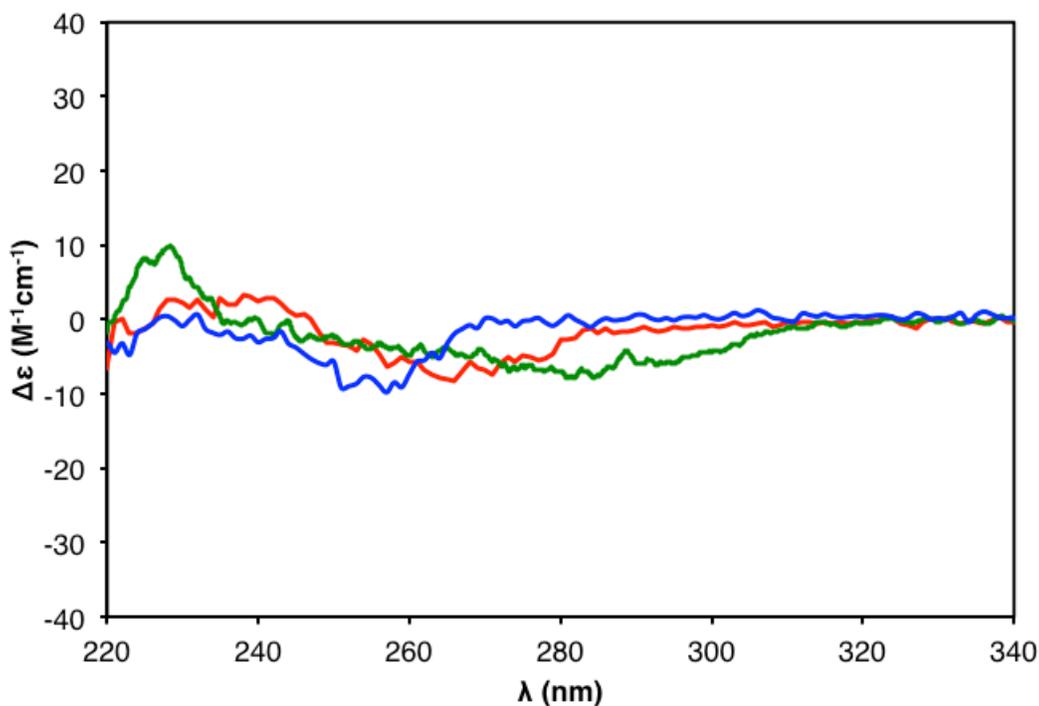
Adenine tetrads in solution stabilized as part of adjacent G-quadruplex structures have been proposed by a only a few groups with Patel *et al.'s* NMR studies being the most convincing.[105,106]

Methanolic CD spectra of the remaining unblocked pendant functionalized conjugates **3a** and **3c-d** with guanine, cytosine, and thymine bases, were compared to aqueous CD spectra and to CD spectra of blocked analogues **9a** and **9c**. Guanine **3a** displayed pronounced magnitude enhancement and wavelength maxima shifts for both positive and negative CD bands in methanol as compared to water, and this enhancement

Figure 4.6   CD spectra of guanine conjugate **3a** in methanol (red) and in 10 mM Tris-borate

pH 8 buffer (blue) and of blocked guanine conjugate **9a** in methanol (green) and

in 10 mM Tris-borate pH 8 buffer (orange).
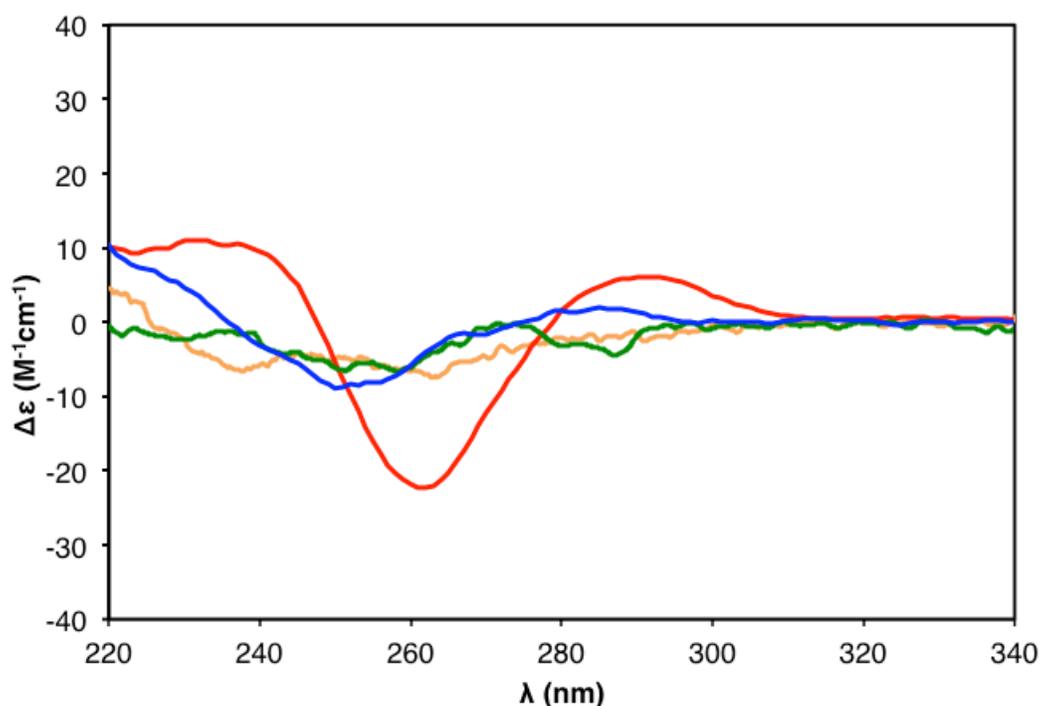
was not observed for blocked guanine conjugate **9a** in methanol or in water (Figure 4.6).

Interestingly, the positive CD at 290 nm and negative CD at 260 nm displayed by

guanine **3a** in methanol reproduced the antiparallel G-quadruplex CD signature. These

Cotton effects constitute significant wavelength and magnitude shifts relative to free

guanosine nucleoside.[150] Cytosine conjugate **3c** showed a change in magnitude for its

positive band in methanol in relation to the spectrum recorded in water (Figure 4.7).

Differences in magnitude and in CD maxima wavelengths were demonstrated for blocked

cytosine conjugate **9c**. Thymine

Figure 4.7   CD spectra of cytosine conjugate **3c** in methanol (red) and in 10 mM Tris-borate

pH 8 buffer (blue) and of blocked cytosine conjugate **9c** in methanol (green).
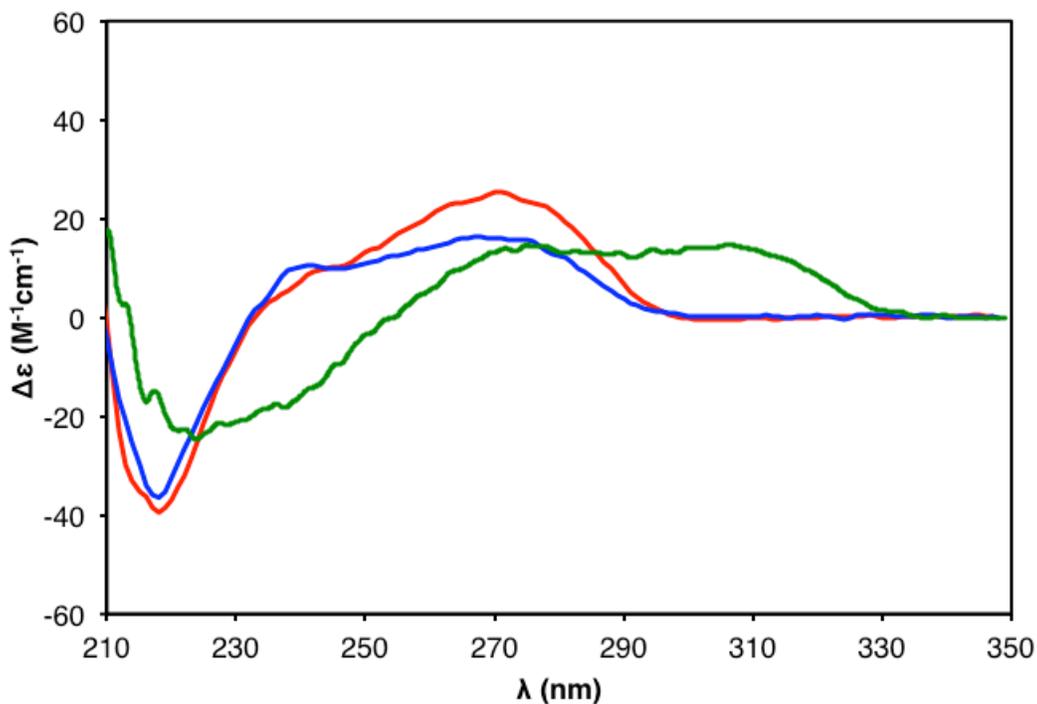
conjugate **3d**, when dissolved in methanol, exhibited enhancement in its respective negative and positive CD bands situated at approximately 220 and 270 nm (Figure 4.8).

Observation of CD enhancement suggests that a self-assembly process is operating but does not constitute proof of tetrad assembly. Enhancements in CD structure were most apparent in methanolic solutions of guanine **3a** and thymine **3d**, and the potential template-assembled behaviour of these conjugates was further characterized by NMR spectroscopy.

Figure 4.8   CD spectra of thymine conjugate **3d** in methanol (red) and in 10 mM Tris-borate

pH 8 buffer (blue).
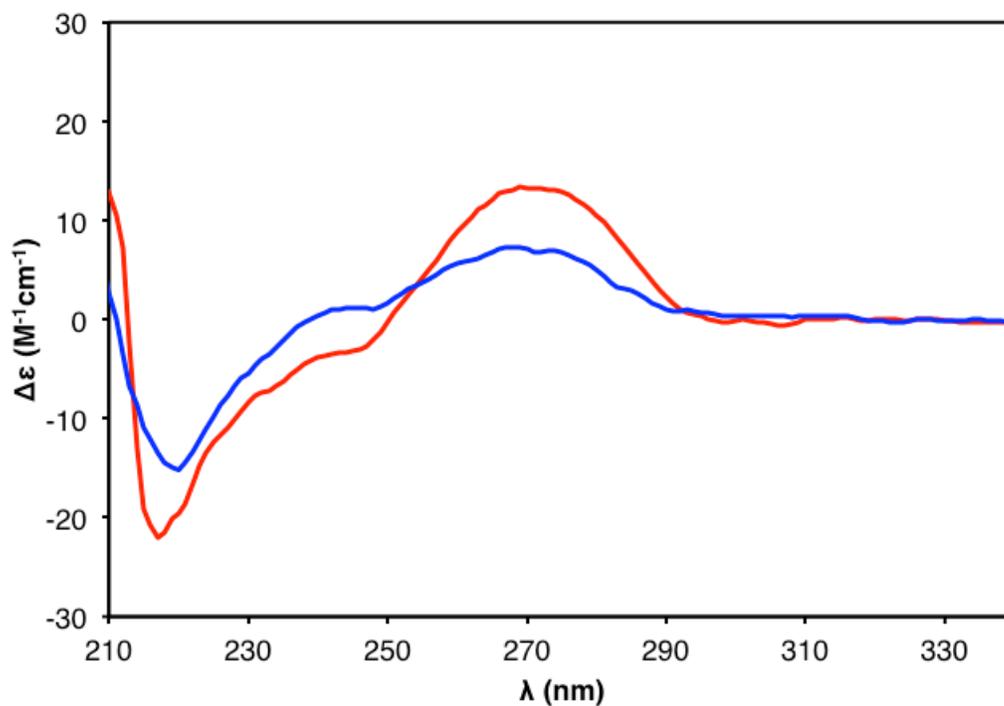
4.3   NMR Spectroscopy of Phosphate-Linked Cavitand-Nucleotide Conjugates in Methanol

The imino proton region in the proton NMR spectrum provides a convenient diagnostic window for strand association and folding in nucleic acids. For example, in duplex DNA, the imino proton resonances of thymine and guanine in hydrogen bonded Watson-Crick G•C and A•T base pairs are typically found in the 12 to 14 ppm range, whereas intitial NMR studies on G-quadruplex DNA, showed imino resonances to be

relatively upfield with chemical shifts between 10.5 and 12.5 ppm.[15,153] Appearance of these downfield imino resonances in buffer is indicative of hydrogen bonded base pairing in nucleic acid structure. Chemical exchange between hydrogen bonded imino protons and solvent molecules is believed to proceed through an opening of the base pair, and catalysis for this exchange process by nitrogen-containing general bases results in exchange line broadening.[154,155] Because quaternary ammoniums do not possess lone pairs and are thus not Lewis bases, preparation of phosphate-linked cavitand-nucleotide conjugates as quaternary ammonium salts was advantageous for studying the dynamics of the imino proton by NMR spectroscopy. Indeed, imino resonances for guanine conjugate **3a** and thymine conjugate **3d** were resolved as lone singlets in methanol solvent in the 11-12 ppm region (Figure 4.9). Resolution of the imino resonance was improved for

Figure 4.9 ^{1}H NMR spectra in 9:1 CH$_3$OH / CD$_3$OD at 400 MHz for a) guanine conjugate **3a** at -25 °C, b) thymine conjugate **3d** at 25 °C, and c) 5'-TBS-2'-deoxythymine at -25 °C.
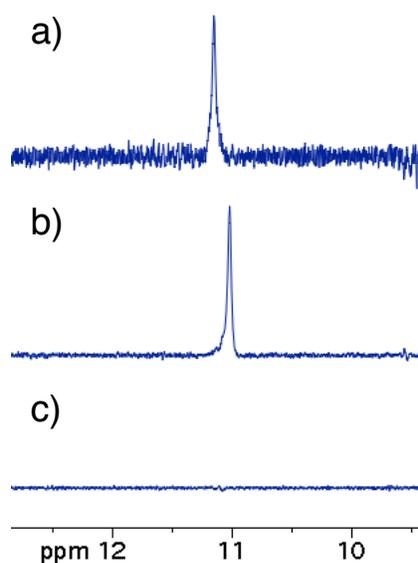
guanine conjugate **3a** by cooling to -25 °C. In comparison, the imino resonance of free nucleoside control compound, 5' TBS protected thymidine, was not observed in methanol solvent even at low temperature. Thus, the fact that this resonance is observed and is present in the characteristic downfield region is highly suggestive that guanine and thymine bases of conjugates **3a** and **3d** are taking part in hydrogen bonded base pairing with exclusion of solvent molecules.

Examination of the ^{1}H NMR spectra for all phosphate-linked tetra-substituted nucleotide conjugates **3a-d** and **9a-c** in methanol reveals that each chemical shift nonequivalent resonance of the nucleoside or cavitand is found to belong to one set of signals (Chapter 2, Section 2.5-6). For example, only one signal is seen for the cavitand aryl proton, the nucleoside CH-1' proton, or the H-8 proton in guanine conjugate **3a** despite four-fold nucleotide conjugation to the cavitand template. That all conjugate nucleotides exist in identical chemical environments is a requirement of C_4 symmetry or equivalence through a rapid process. Nucleobase quartets share a common C_4 symmetry axis directed through the center of the quartet orthogonal to the quartet plane, and this symmetry element would be retained in a unimolecular template-assembled nucleobase quartet (Figure 4.10). This NMR symmetry is characteristic of known unimolecular template-assembled quartets.[91,103]

Conformational analysis of G-quadruplex DNA is routinely performed by nuclear Overhauser effect spectroscopy (NOESY) NMR.[156] In the NOE experiment, information about the spatial proximity between different nuclei interacting through dipolar coupling is provided. The intensity of a dipolar interaction between two different nuclei is inversely dependent on the internuclear distance to the sixth power.

Figure 4.10   A C_4 axis of symmetry is preserved between a) a single nucleobase quartet and

b) a template-assembled single nucleobase quartet.
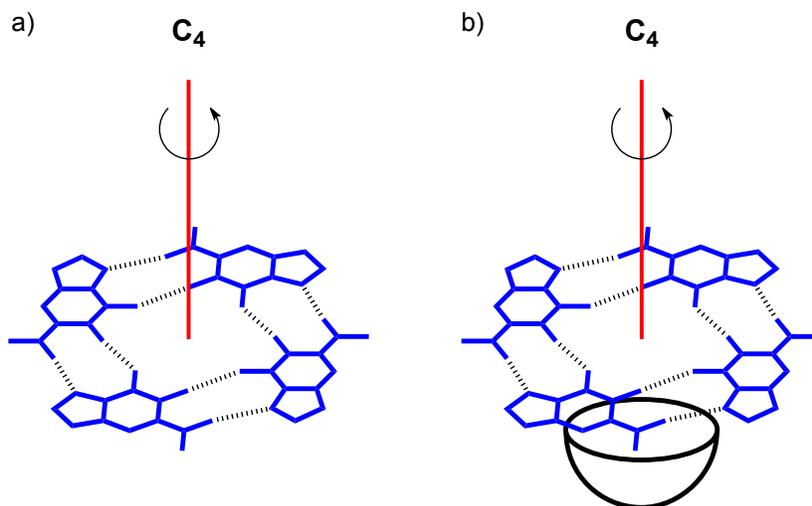
The presence of a *syn* glycosidic bond conformation is detected using a 2D NOESY protocol when a NOE cross peak between the CH-1' resonance of the ribose sugar and the H-8 resonance of the guanine base exists (Figure 4.11).[15] A NOE correlation indicating a *syn* conformation was found for guanine conjugate **3a** in deuterated methanol at -25 °C. Previously reported template-assembled synthetic quartets have existed as *syn* conformers.[91,103,140]

For large molecules, NOE cross peaks may not be observable in the NOESY NMR experiment. The rotating frame Overhauser effect spectroscopy (ROESY) experiment alleviates this complication by improving sensitivity through the provision of positively phased cross peaks.[157] Comprehensive structural assignment of thymine conjugate **3d** in pure methanol

Figure 4.11 **a)** Glycosidic bond *syn* conformation gives rise to NOE correlation between guanine base and sugar protons. **b)** 2D NOESY spectrum in CD$_3$OD at 400 MHz for guanine conjugate **3a** at -25 °C indicating *syn* conformation. Mixing time was 800 ms.
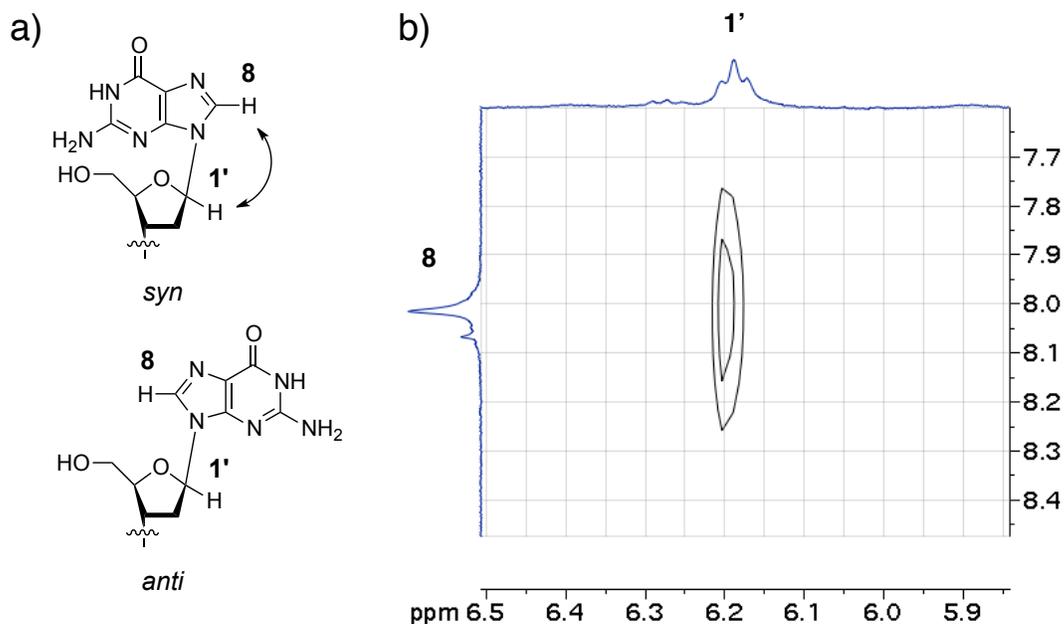
was pursued by a 2D ROESY experiment with simultaneous methanol signal suppression (Figure 4.12). Observed ROE cross peaks between the imino and 5' methyl resonances and between the H-6 and CH-1' resonances support the existence of a planar template-assembled thymine tetrad with *syn* glycosidic conformers. Full characterization of a thymine tetrad has only been reported recently for a cation stabilized T-tetrad capped G-quadruplex DNA.[158]

Figure 4.12   Observed ROE correlations in 2D ROESY NMR experiment in 9:1 CH$_3$OH /

CD$_3$OD at 400 MHz for thymine conjugate **3d** at -15 °C. A ROESY spin-lock pulse

of 800 ms was used.
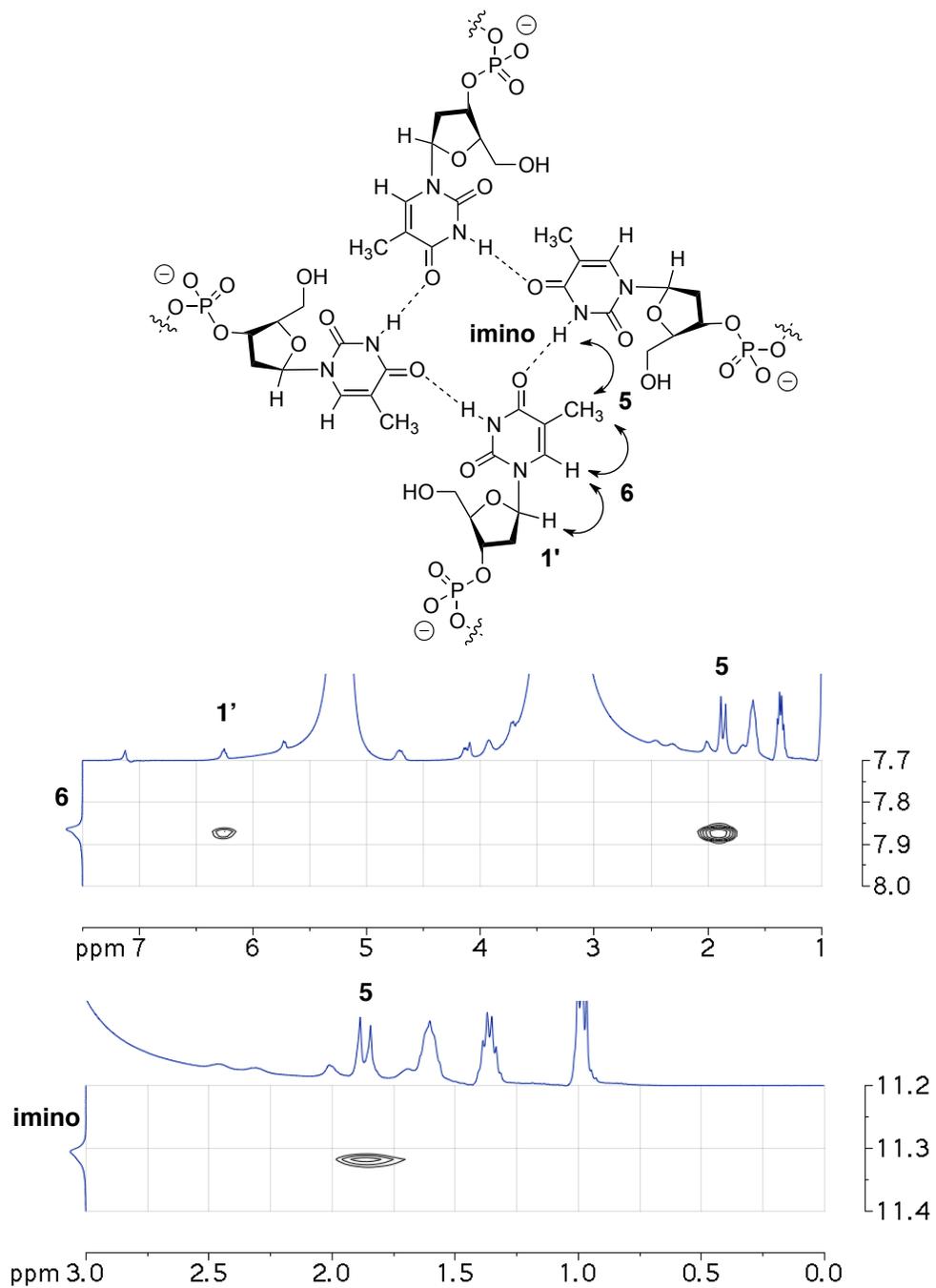
4.4 Diffusion Ordered NMR Spectroscopy of Phosphate-Linked Cavitand-Nucleotide Conjugates in Methanol

Diffusion ordered NMR spectroscopy or DOSY is an important type of pulsed field gradient NMR and has become increasingly utilized for the evaluation of the size and shape of molecular aggregates in solution by means of measuring translational diffusion coefficients.[159,160] Applications of DOSY in the study of nucleobase assemblies include the determination of precise lengths for cylindrical G-quartet stacks of 5'-GMP in basic solution, of stoichiometries of several cation-templated lipophilic guanosine-derived assemblies, and of guanine and uracil dimerization for template-assembled synthetic quartets in chloroform.[92,103,161,162]

The translational diffusion of a spherical molecule in a medium is defined by the Stokes-Einstein equation:

$$D = \frac{kT}{6\pi\eta r_s}$$

Where D is the diffusion coefficient, k is the Boltzmann constant, T is the temperature, η is the viscosity of the medium, and r_s is the hydrodynamic radius of the molecule. For two spherical molecules with different hydrodynamic radii existing in the same environment at constant temperature, it can be seen clearly from this expression that the smaller molecule will diffuse at a faster rate and thus be quantified by a greater diffusion coefficient.

The simplest pulse sequence that can be used to measure diffusion is the pulsed gradient spin echo experiment first developed by Stejskal and Tanner in 1965.[159,164] Application of two magnetic field gradient pulses separated by a time interval allows labeling of nuclear positions in a sample and an extraction of a diffusion coefficient. An exponential decay of the signal intensity, I is observed during the course of the experiment and is described by the following equation:

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)}$$

Where I_0 is the unperturbed signal intensity, γ is the gyromagnetic ratio for the observed nucleus, g is the gradient strength, δ is the length of the gradient pulse, and Δ is the diffusion time. The decay of the signal intensity as a function of gradient strength is measured at constant temperature and at a defined solute concentration with optimized values for the gradient pulse length and diffusion time, and this allows a determination of a diffusion coefficient for the solute.

Circular dichroism and NMR studies described in previous sections have provided strong evidence that guanine and thymine conjugates **3a** and **3d** exist as hydrogen bonded self-assemblies with defined conformation in methanol solvent. However, it is not clear from these studies alone whether the assemblies are intermolecular or intramolecular by nature. Using 2D DOSY methods, diffusion coefficients were measured for phosphate-linked cavitand-nucleotide conjugates **3a** and **3c-d** and the blocked analogues **9a** and **9c** in order to ascertain whether intermolecular self-association was taking place for samples of these compounds (Figure 4.13, Table 4.1). The base-blocked analogue was chosen as an ideal control compound for DOSY experiments, since its hydrogen bonding capacity

Figure 4.13   Data points from a 2D DOSY NMR experiment in CD_3OD at 400 MHz for a 7.1

mg/mL solution thymine conjugate **3d** at 25 °C show the decay of signal intensity

as a function of gradient strength. Curve fitting (blue) gave a diffusion coefficient

of 2.72×10^{-10} m^2s^{-1}.
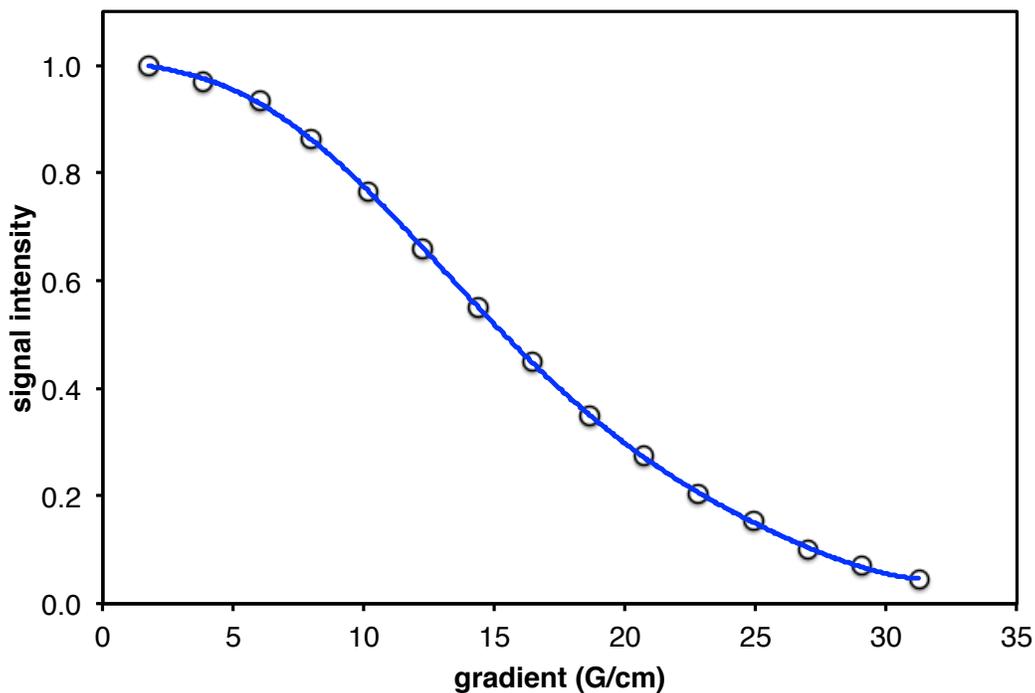
has been removed. Thus, nucleobase mediated intermolecular hydrogen bonding between

different molecules in solution should be excluded for a sample of the control compound.

If identical diffusion coefficients are measured between the unblocked compound and the

blocked control compound, for example between conjugates **3a** and **9a**, then it can be

assumed that both compounds are associated to the same extent in solution. Furthermore,

nucleobase assembly in the unblocked compound, in this example **3a**, must be of a single

template-assembled guanine quartet.

Table 4.1   Diffusion coefficients, *D* for 7.1 mg/mL methanolic solutions of phosphate-linked

cavitand-nucleotide conjugates at 25 °C.

compound	base	D / 10^{-10} m²s⁻¹
3a	G	2.54
9a	G^{Ibu}	2.56
3c	C	2.67
9c	C^{Bz}	2.64
3d	T	2.72

The diffusion coefficients for the unblocked and isobutyryl protected guanine based conjugates **3a** and **9a** were determined to be 2.54 and 2.56 × 10^{-10} m²s⁻¹ respectively, and these differ by less than 1 %. Similarly, a 1 % discrepancy is seen for cytosine derived conjugates, and here diffusion coefficients were found to be 2.67 × 10^{-10} m²s⁻¹ and 2.64 × 10^{-10} m²s⁻¹ for **3c** and **9c**. Dimerization via nucleobase elements represents the first potential level of self-association, and a dimer species would be expected to display a diffusion coefficient of approximately 79 % of the control compound according to an inverse cube root dependence between the ratios of diffusion coefficients, *D* and molecular weights, *M* for two different spherical molecules:[92,103,165]

$$\frac{D_1}{D_2} = \sqrt[3]{\frac{M_2}{M_1}}$$

For thymine conjugate **3d**, no blocked analogue was synthesized. However, its diffusion coefficient of 2.72 × 10^{-10} m²s⁻¹ is 97-98 % of fellow pyrimidine compounds **3c** and **9c**,

and it is reasonable to conclude that thymine conjugate **3d** exists in an identical extent of association as **3c** and **9c**. The simplest conclusion to be made from the DOSY data presented here is that all guanine, cytosine, and thymine derived cavitand-nucleotide conjugates exist as unimolecular species in methanol solvent. Therefore, unimolecular guanine and thymine template-assembled nucleotide quartets have been synthesized (Figure 4.14).

Figure 4.14   Cartoon representations of unimolecular template-assembled nucleobase quartets by guanine conjugate **3a** (blue) and thymine conjugate **3d** (red). The cavitand (black) has its concave surface directed away from the quartet.
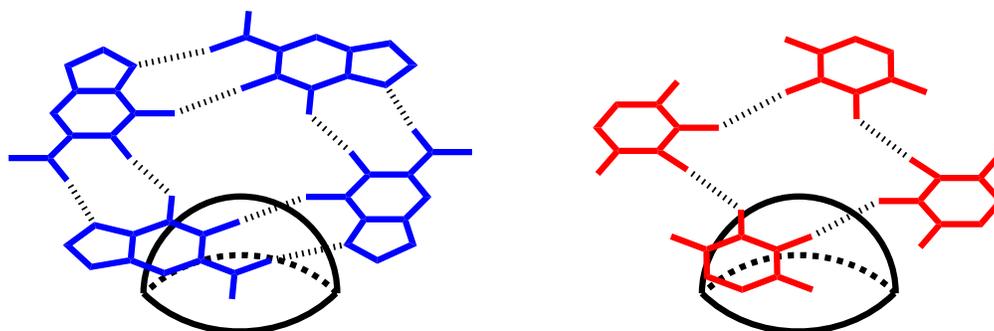
4.5   The Role of Cations in Template-Assembly

It is well understood that assembly of G-rich oligonucleotides and guanosine monophosphates in water and of guanosine nucleosides in lypophilic systems into G-quartet and G-quadruplex structures is highly dependent on metal cation stabilization.[166]

Reports exist for G-quartets forming from free or template bound nucleosides in the absence of cations in organic solvent and in the solid-state, but these are relatively rare.[91,167,168] Therefore, an analysis of potential metal contaminants accompanying the production of phosphate-linked cavitand-nucleotide conjugates by newly developed phosphite triester and phosphoramidite synthetic methodologies was carried out (Chapter 2, Section 2.5). Inductively coupled plasma mass spectrometry (ICP-MS) on selected samples of conjugate was performed (see Section 4.6). Sodium is persistent in the environment, and for example, the sodium atom was found to make up only 6.1 mol % for a sample of thymine conjugate **3d**. Additional G-quartet stabilizing metal contaminants such as potassium were not assumed to be significant based on their absence in the synthetic preparation and in the purification methods employed. Stoichiometrically, there is a four to one molar ratio between guanine base and G-quartet stabilizing cation, and large molar excesses of cation are normally required to induce guanine assembly in aqueous media. The synthetic preparations of phosphate-linked cavitand-nucleotide conjugates here were shown to produce less than 10 mol % sodium contamination, and self-assembly behavior by any of these conjugates can be regarded as being metal cation independent.

4.6 Experimental Procedures

Circular dichroism spectroscopy was performed using a Jasco J-815 CD spectrometer. Samples were prepared in the 0.01 to 0.02 mM concentration range in 10 mM Tris-borate pH 8 buffer or methanol in a 1 cm cell. Millipore water or spectroscopic grade methanol was used in the preparation of these solutions. The ROESY NMR experiment was performed by incorporating a water suppression enhanced through T_1 effects (WET) pulse sequence into a standard Bruker 2D ROESY program to suppress both methanol signals. A standard Bruker program was used for 2D DOSY experiments. Big delta and little delta values were optimized to 100 ms and 4 ms respectively for all experiments. Gradient strength was increased over 16 steps from 2 % to 95 %. Diffusion coefficients were obtained from the signal intensity decay of the ribose sugar 1' proton as a function of gradient strength after curve fitting using the XWINNMR SimFit algorithm. Sodium analysis was performed on a Thermo Fisher Element XR ICP-MS. Samples of 1-2 mg of conjugate were dissolved in 1% trace element nitric acid and diluted accordingly. Concentrations of sodium in the conjugate samples were determined using standard addition after calibration with sodium-23 and indium-115 internal standards and subtraction from a control sample.

Chapter 5: Template-Assembled Isolation of a Guanine Tetrad in Water and its Application as an Artificial Receptor for G-quadruplex Binding Ligands

5.1 Synopsis

A phosphate-linked cavitand-nucleotide conjugate functionalized at the rim position was evidenced to template-assemble a guanine tetrad in water. Support for this conclusion was provided for by CD, ^{1}H NMR, DOSY, and NOESY NMR experimentation. A strong induced CD spectrum was observed relative to an analogue conjugate blocked at the base position. Analysis of the ^{1}H NMR spectrum in water showed an imino proton signal characteristic for guanine-guanine Hoogsteen hydrogen bonding in G-quadruplex DNA and also allowed the assignment of the amino resonance. A *syn* glycosidic bond conformation was confirmed by a 2D NOESY experiment. Diffusion coefficients determined by 2D DOSY were shown to be indistinguishable between the free base and the blocked conjugate. Sodium content in a sample of conjugate was shown to be less than 10 mol % by ICP-MS. However, an investigation into the role that cations may play in stabilizing or binding to this template-assembled G-tetrad failed due to ambiguous CD and NMR data. A cation-free unimolecular template-assembled G-tetrad structure with one face exposed to solvent was proposed, and this

represents the first known observation of a single unstacked G-tetrad in aqueous solution. Next, the potential for this cavitand-nucleotide conjugate to act as an artificial receptor for small molecules known to inhibit telomerase and bind to G-quadruplex DNA was probed. Mixtures of conjugate and TMPyP4 ligand or conjugate and PIPER ligand displayed characteristic CD structure consistent with a G-tetrad-ligand interaction. Finally, a binding constant of 8.5×10^5 M^{-1} and 1:1 binding stoichiometry was measured by Scatchard and Job plots using UV / VIS spectrophotometry for TMPyP4.

5.2   Template-Assembled Isolation of a Guanine Tetrad in Water

An investigation of rim functionalized phosphate-linked cavitand-nucleotide conjugate **7**, synthesized in Chapter 2, to template-assemble a guanine tetrad in water was carried out according to the experimental techniques, namely CD, ^{1}H NMR, NOESY, and DOSY, outlined previously in the elucidation of template-assembled nucleobase tetrads in methanol (Figure 5.1; see Sections 4.2-6). A more rigorous theoretical discussion can be found there, and the arguments used therein to justify the possibility of these methanolic nucleobase quartets should be equally applicable to the situation of water-solubility described here. Cavitand-nucleotide conjugate **9a**, which is blocked at its base exocyclic amino position with isobutyryl protection, has restricted hydrogen bonded self-assembly capability and was selected as a comparison compound.

118

Figure 5.1   Phosphate-linked guanine base cavitand-nucleotide conjugates with unblocked

residues at the rim position and blocked residues at the pendant position.

7

9a
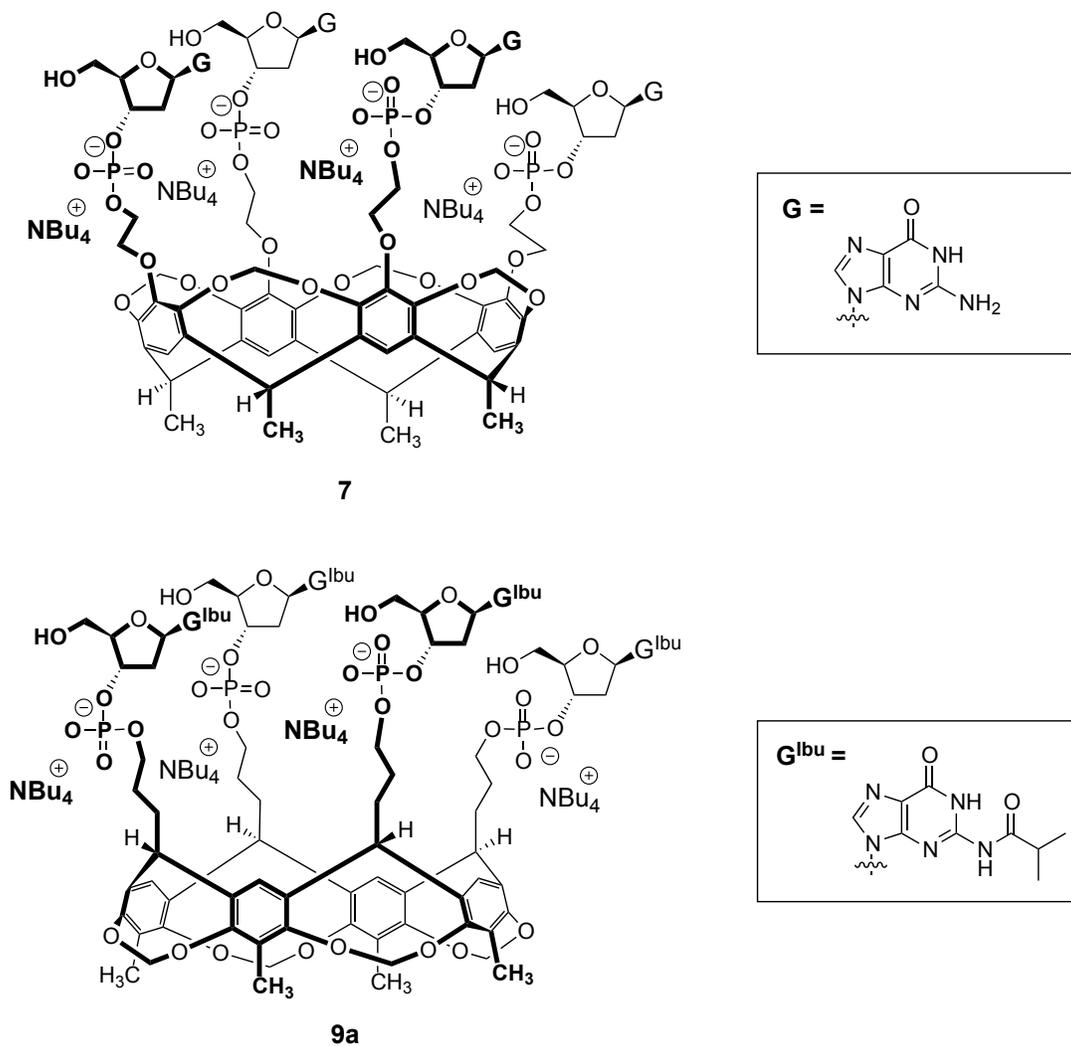
Circular dichroism spectra of guanine conjugate **7** and of blocked conjugate **9a** in

methanol and buffer are given (Figure 5.2). Guanine conjugate **7** displays CD structure in

methanol in accordance with pendant functionalized cavitand-nucleotide conjugates (see

Figure 5.2 CD spectra of 0.02 mM guanine conjugate **7** in methanol (red) and in 10 mM Tris-borate pH 8 buffer (blue) and of 0.02 mM blocked guanine **9a** in methanol (orange) and in 10 mM Tris-borate pH 8 buffer (green).
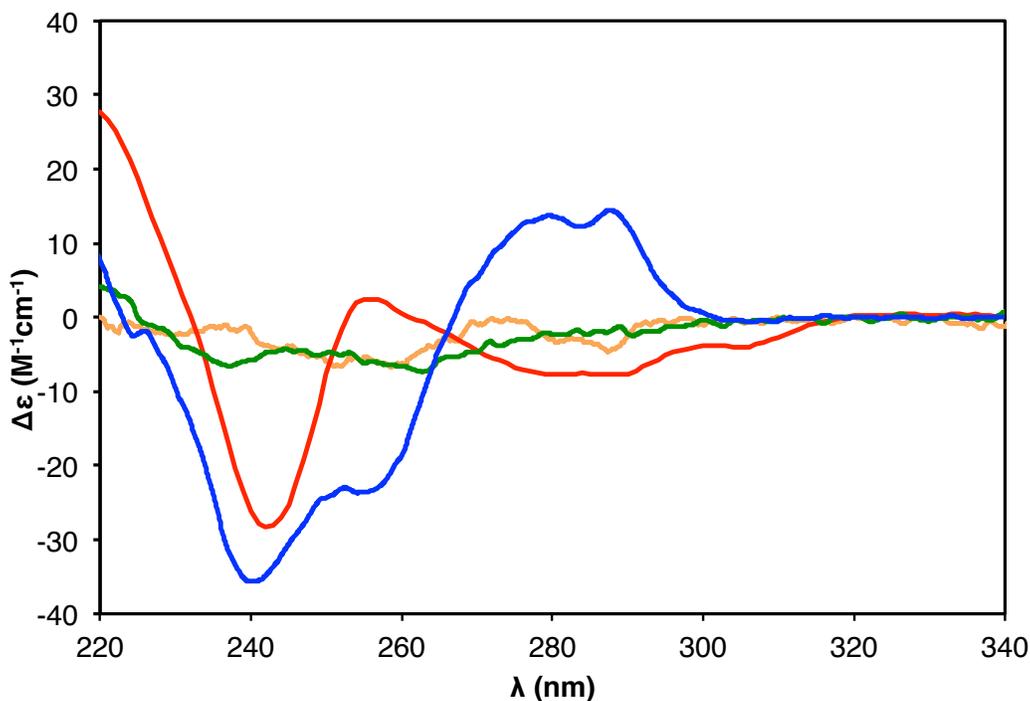
Section 4.2). Attention should be drawn to the dramatic CD enhancement displayed for guanine conjugate **7** in water with respect to control compound **9a**. Strong CD induction in water was not found previously for either triazole-linked cavitand-guanosine constructs or a pendant-functionalized phosphate-linked cavitand-guanylate conjugate and thus, is a unique property to conjugate **7** (see Sections 3.4 and 4.2). Induction of CD is explained through exciton coupling between chromophores orientated both chirally and with close spatial proximity and provides evidence for helical self-assembly of guanine bases.[142] Electronic transitions of the π-π* type in guanine give rise to two well resolved

120

absorption bands in the UV region at approximately 250 and 279 nm. A wavelength of maximum absorption in the absorption spectrum for an exciton coupled chromophore typically corresponds to zero CD intensity with positive and negative CD on either side. Empirically, parallel G-quadruplex DNA shows a positive exciton coupling with a negative band at 240 nm and a positive band at 260 nm. In contrast, antiparallel G-quadruplex DNA is characterized by a 260 nm negative band and a 290 nm positive band. Two overlapping negative bands close to 240 and 255 nm and two overlapping positive bands at approximately 280 and 290 nm can be seen in the CD spectrum of conjugate **7** in water, and correspond to mixed empirical features of both parallel and antiparallel G-quadruplex DNA. However, care should be taken not to infer topological structure, for example glycosidic torsion angles, in this case, because an empirical interpretation of the CD of G-quadruplex DNA is limited to stranded systems of oligonucleotides. It is conceivable that the CD spectrum of **7** in water originates from two dominant overlapping positive exciton couplets centered near to the adsorption maxima of 250 and 279 nm for free guanine (Figure 5.3).

The ^{1}H NMR spectrum for a solution of conjugate **7** in water exhibits a downfield guanine imino proton resonance that is observable at room temperature in the ppm range characteristic of G-quadruplex DNA and is resolved at lowered temperatures (Figure 5.4).[15] Hoogsteen guanine-guanine hydrogen bonded base pairing slows chemical exchange between imino protons and water solvent molecules facilitating their observation in the NMR spectrum. A new broadened signal in the water suppressed ^{1}H NMR spectrum at approximately 6.3 ppm also resolved by variable temperature was

Figure 5.3   Correlation of the Absorbance and CD spectra of guanine conjugate **7** in 10

mM Tris-borate pH 8 buffer.
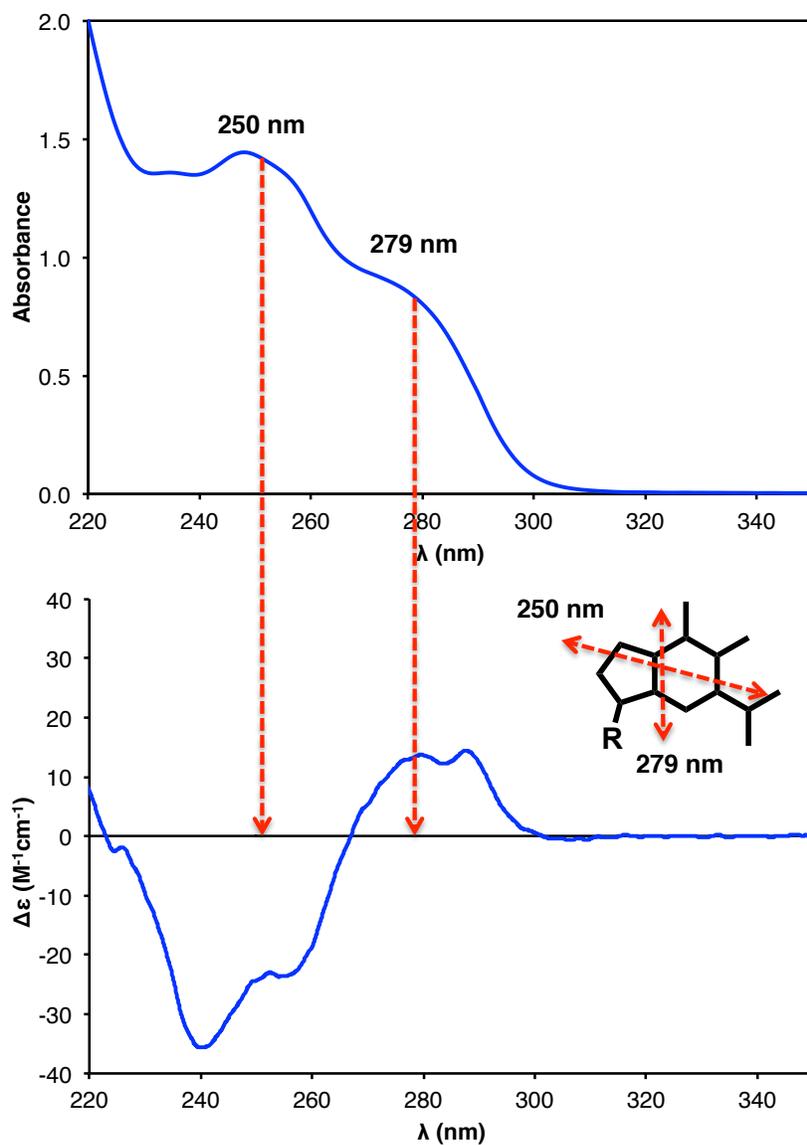
assigned to the amino protons of the guanine base. Appearance of this signal as a single

broadened peak indicates that the amino protons are chemically equivalent through fast

bond rotation on the NMR time scale.[169] Amino and imino protons localized to the

terminal base pairing regions are often not seen in the NMR spectrum due to rapid exchange with solvent through strand fraying processes.[170] It is noteworthy that the exchangeable base protons of conjugate **7** are kinetically stabilized despite a potential that one face of **7**'s guanine assembly is exposed to solvent molecules. The C_4 symmetry requirement for a single template assembled guanine tetrad arrangement is satisfied by the appearance of one set of signals for all four nucleotides and the cavitand in the NMR for conjugate **7**.

Figure 5.4 ^{1}H NMR spectra of guanine **7** in 9:1 H_2O / D_2O at 400 MHz for the a) imino region
and b) amino (black arrow) region with suppression of water by presaturation.
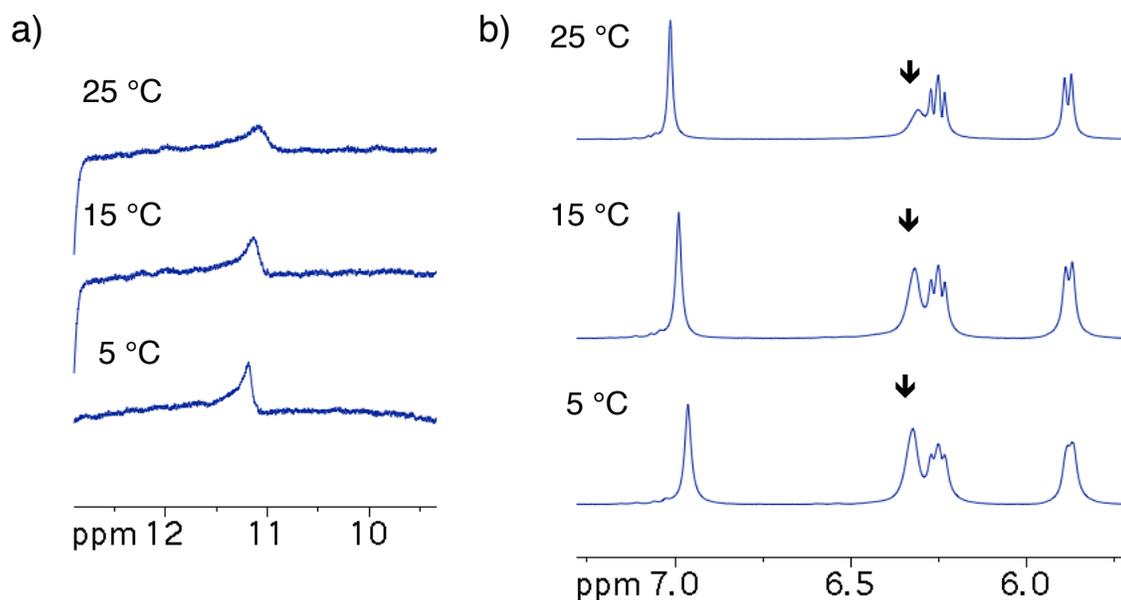
Two-dimensional NMR methods in deuterated water were utilized to gain information regarding the conformation and the extent of self-association for guanine

conjugate **7**. A *syn* glycosidic torsion angle is evidenced by an NOE correlation between the CH-1' ribose proton at 6.2 ppm and the H-8 base proton at 8.0 ppm in the 2D NOESY

Figure 5.5   a) Proposed G-tetrad assembly with *syn* glycosidic torsion angles for conjugate **7** indicating a NOE interaction between sugar and base protons. b) 2D NOESY spectrum of **7** in H_2O / D_2O at 400 MHz and 5 °C with a mixing time of 800 ms.
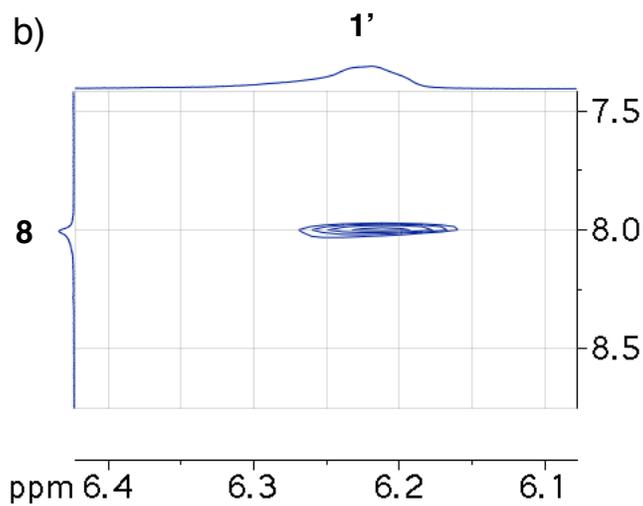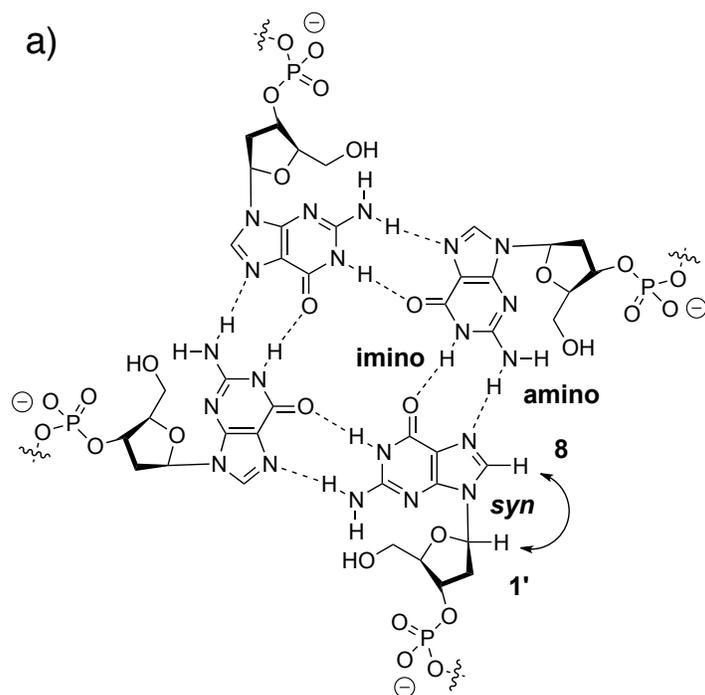
NMR spectrum (Figure 5.5). Diffusion coefficients were measured by 2D DOSY experiments at 25 °C for 7.1 mg/mL solutions and were determined to be 1.43×10^{-10} m^2s^{-1} for guanine **7** and 1.45×10^{-10} m^2s^{-1} for blocked guanine **9a**. The diffusion coefficients reported here match to within 98.7 % providing evidence that conjugate **7** is not associated in solution to any significant extent and is template-assembling a single intramolecular guanine tetrad in water.

Sodium content in a sample of guanine conjugate **7** was determined to be 8.3 mol % by ICP-MS analysis. Interestingly, CD spectra in buffer were unchanged in the presence of 50 mM NaCl or KCl (Figure 5.6). Furthermore, CD data for alkali and alkaline earth metal chlorides presented the same result (data not shown). Proton NMR spectra obtained for **7** dissolved in aqueous solutions of sodium or potassium chlorides or phosphates displayed considerable line broadening that rendered spectra uninterpretable (data not shown). Thus, it can be concluded that conjugate **7** templates a guanine quartet in the absence of cation stabilization. However, the role if any that cations such as sodium or potassium play in binding to this unimolecular template-assembled guanine quartet remains uncertain.

Figure 5.6   CD spectra of 0.02 mM guanine conjugate **7** in 10 mM Tris-borate pH 8 buffer

(blue dashed) and in the presence of 50 mM NaCl (orange) or KCl (green).
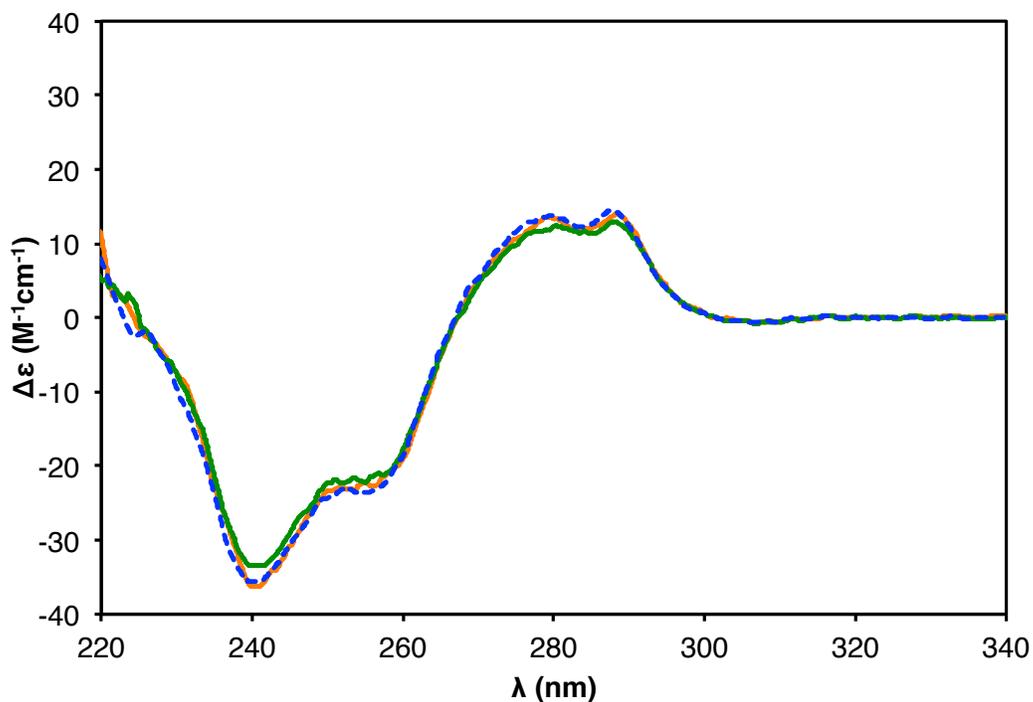
5.3   Structural Significance of an Isolated Template-Assembled Guanine Tetrad in Water

Experimental observations supporting the isolation of a single cation-free self-assembled guanine tetrad in water through the use of covalent template organization of four attached nucleotides have been described thus far. This work represents the first known example of an individual G-tetrad in aqueous solution. It is well understood that

the assembly of guanine moieties through Hoogsteen hydrogen bonding into the canonical cyclic tetramer structure is intrinsically unstable in water. Reported structures for guanine assemblies involving oligonucleotides and nucleotides share two defining qualities in that G-tetrads are always found to be associated in stacks and that these assemblies depend on sodium or potassium stabilization.[171] Template-assembly of G-quadruplexes has become a novel strategy for molecular preorganization, and several precedents exist in addition to the work presented here. Oligonucleotides have been organized into parallel G-quadruplex DNA through the use of tetra-end-linked and peptidic templates.[94,95,97,98] As well, the same peptidic template has been functionalized with four guanosine residues, but was shown to direct a dimeric species in water.[140]

In the context of the thesis aims outlined previously, certain questions remain to be addressed (see Section 1.8). How closely does the guanine tetrad structure of this unimolecular template-assembled cavitand-nucleotide construct resemble human telomeric G-quadruplex DNA? Can this exposed G-quartet face act as a receptor for G-quadruplex binding ligands? Can a template-assembled synthetic G-quartet have practical applications, for example, in discovering potential anticancer therapeutics?

Examination of all reported human telomeric DNA G-quadruplex structures both in solution and in the solid-state demonstrates that folding of oligonucleotides comprising the TTAGGG repeat is polymorphic by nature. At the present time, the variability between these proposed structures precludes a firm conclusion with respect to the folding topology of human telomere DNA *in vivo*, and it is possible that several of these topologies exist in dynamic equilibrium in the cell. The first known human G-quadruplex solution structure was presented by Wang *et al*. for the 22mer oligonucleotide

127

d[AGGG(TTAGGG)$_3$] in sodium solution.[39] The resultant structure was of a mixed antiparallel / parallel intramolecular G-quadruplex consisting of three G-tetrad stacks connected by lateral and diagonal loops. The glycosidic torsion angles were found to alternate within a strand and within each tetrad such that a *syn / anti / anti / syn* or *anti / syn / syn / anti* pattern was present. Later, Parkinson and coworkers reported that a Wang's 22mer human telomere sequence crystallized under physiological potassium concentrations into a self-folded parallel G-quadruplex with propeller loops and all *anti* conformations.[40] An intermolecular dimeric G-quadruplex crystal structure with similar topologies otherwise for a shortened 12mer telomeric sequence in potassium was also included in this report. Potassium concentrations are greater at the intracellular level than sodium concentrations, and this fact was used to garner support for greater physiological relevance to potassium human G-quadruplex structures. However, an even newer G-quadruplex polymorphism was provided after solving the potassium solution structure for the 22mer human telomere DNA.[172-174] This later intramolecular G-quadruplex was found to have mixed antiparallel / parallel strands with different loop regions and with a novel [3+1] glycosidic torsion angle motif. Two of the three stacked G-tetrads were *syn / anti / anti / anti* and the remaining tetrad was *anti / syn / syn / syn*. DNA G-quadruplex formed from sequences found outside the telomere regions in the human genome, such as in oncogene promoters, has also gained interest as a target for anticancer drugs, and these G-quadruplexes show polymorphisms distinct from those of telomeres.[175]

The guanine tetrad templated by construct **7** shares structural similarities and differences to the tetrads formed within known human G-quadruplex DNA structures. Compound **7** conserves the chemical connectivities and composition of DNA, and It

contains 2'-deoxynucleotides with the correct 3' phosphate ester linkage. In contrast, the all *syn* G-tetrad observed for **7** is unique to the glycosidic torsion angle patterns observed for human telomeric DNA. Conjugate **7** is believed to possess an unobstructed G-quartet face that is exposed to solvent molecules. This face may provide a model for the terminal G-tetrad of G-quadruplex structures, which is often involved in end-stacking binding modes with interacting G-quadruplex binding ligands.[53] The ligand binding potential of construct **7** is explored in the following section.

5.4   Application of a Water-Soluble Template-Assembled G-tetrad as an Artificial G-quadruplex Binding Ligand Receptor

Potential interactions in solution between guanine conjugate **7** and two small molecules with well established G-quadruplex DNA binding affinity and telomerase inhibition were probed with CD spectroscopic and UV / Vis spectrophotometric methods (Figure 5.7). The cationic porphyrin and known telomerase inhibitor, TMPyP4, has been demonstrated to interact with a diverse set of G-quadruplex polymorphisms.[66-68,70,72,73,176,177] Circular dichroism spectra for the interaction of TMPyP4 and **7** are shown (Figure 5.8). A strong induced negative CD band in the Soret region was observed for the porphyrin upon mixing with guanine conjugate **7** indicating its interaction.[66] A change in the CD structure in the 220 to 300 nm range supports the notion that TMPyP4 is binding

Figure 5.7   Small molecule G-quadruplex binders and telomerase inhibitors used to

probe the receptor potential of guanine conjugate **7**.

TMPyP4
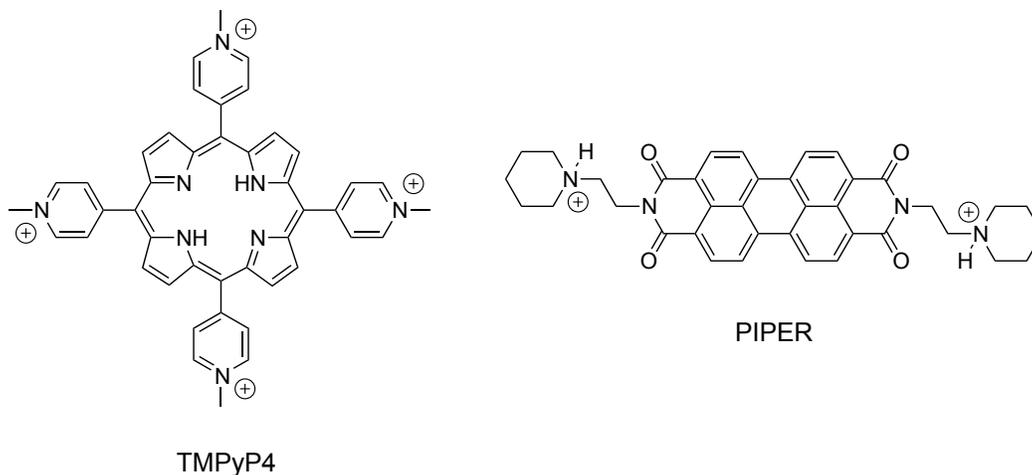
PIPER

to the G-tetrad of **7**. Similarly, the perylene diimide based G-quadruplex stabilizing PIPER ligand exhibited G-tetrad mediated interaction with guanine conjugate **7** (Figure 5.9).[65,178-181] Addition of PIPER ligand resulted in respective broad positive and negative induced CD at approximately 480 and 550 nm together with a change to the characteristic G-tetrad signal.

Quantitative analysis of the binding strength and stoichiometry between guanine conjugate **7** and TMPyP4 ligand was pursued using UV / VIS spectrophotometric methods. Scatchard analysis adapted from duplex DNA-drug binding studies has been successfully applied to TMPyP4 interactions with G-quadruplex DNA in solution.[66,68,182] Free TMPyP4 has a wavelength of maximum absorption situated at 422 nm for its Soret band, and binding of the porphyrin is characterized by hypochromocity and a bathochromic

Figure 5.8 CD spectra in water of guanine conjugate **7** alone (blue), TMPyP4 ligand alone (red), and **7** and TMPyP4 mixed (orange). Concentrations were 0.01 mM.
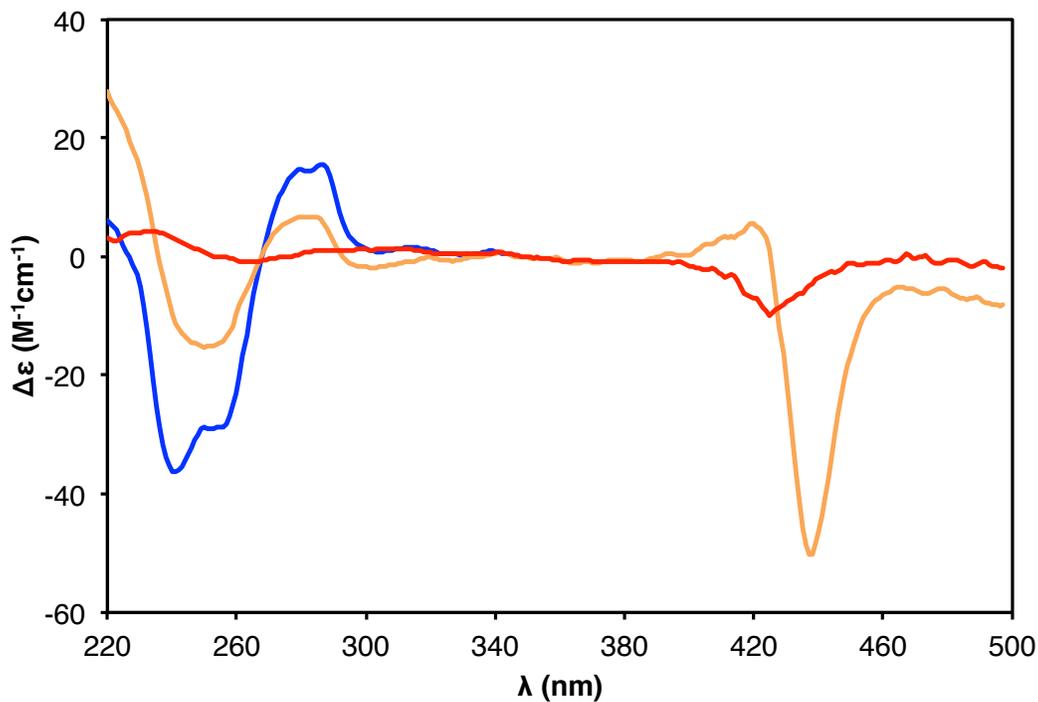
shift. Optical monitoring of binding at 422 nm and fitting to the simple Scatchard equation allows a determination of an equilibrium binding constant, K for the interacting ligand and DNA receptor:

$$r/C_f = K(n - r)$$

where r is the binding ratio between porphyrin and DNA receptor, C_f is the concentration of unbound porphyrin, and n is the number of binding sites per DNA receptor. A binding constant is obtained from the slope of a line fit to a plot of r/C_f versus r. A Scatchard plot

Figure 5.9   CD spectra in water of guanine conjugate **7** alone (blue), PIPER ligand alone

(red), and **7** and PIPER mixed (orange). Concentrations were 0.01 mM.
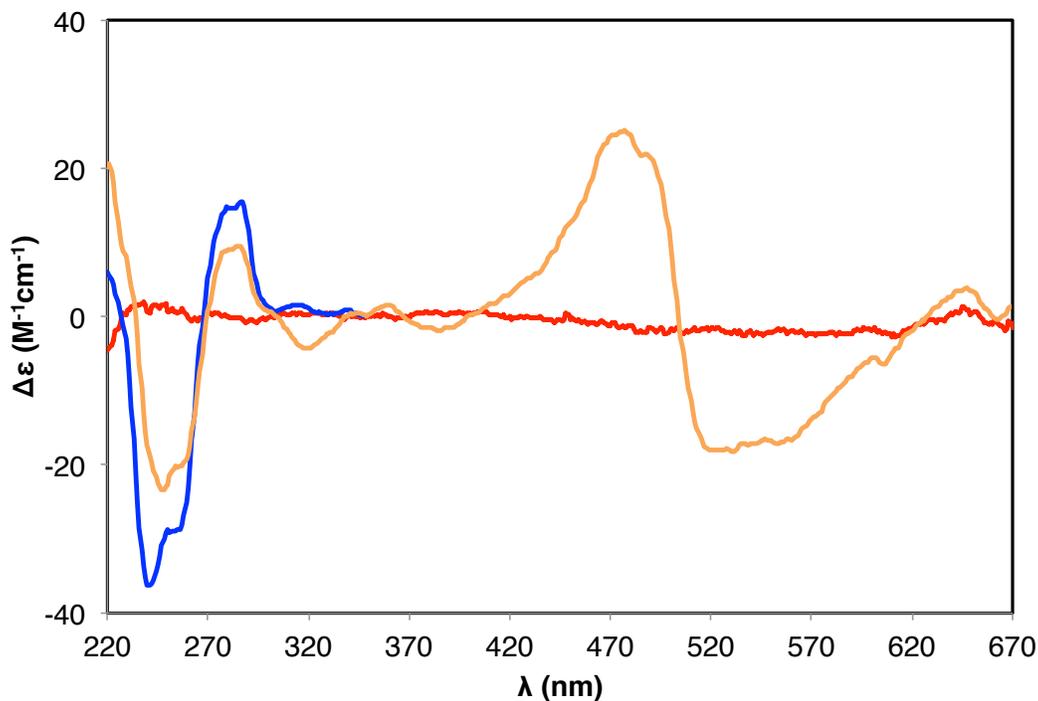
for the binding of TMPyP4 to conjugate **7** gave a binding constant of 8.5×10^5 M^{-1} in accordance with binding constants reported in the 10^4-10^6 M^{-1} range for the interaction with G-quadruplex DNA containing both human and non-human telomere repeat sequences (Figure 5.10).[68]

The method of continuous variation was used in the construction of a Job plot (Figure 5.11).[68,142] An inflection point centered above 0.5 mol fraction of ligand reflects a saturating stoichiometry that is slightly greater than 1:1 for TMPyP4 ligand to conjugate **7**, and this is consistent with a value of $n = 1.4$ obtained from Scatchard's equation. Loss

Figure 5.10   Scatchard plot monitored at 422 nm for the titration of 5 μM TMPyP4 ligand with guanine conjugate **7** in 10 mM Tris-borate pH 8, 0.1 mM EDTA. The slope and y-intercept of the resultant fitted line were -8.5 and 12 respectively. An R^2 value of 0.98 demonstrated good linearity of the data points.
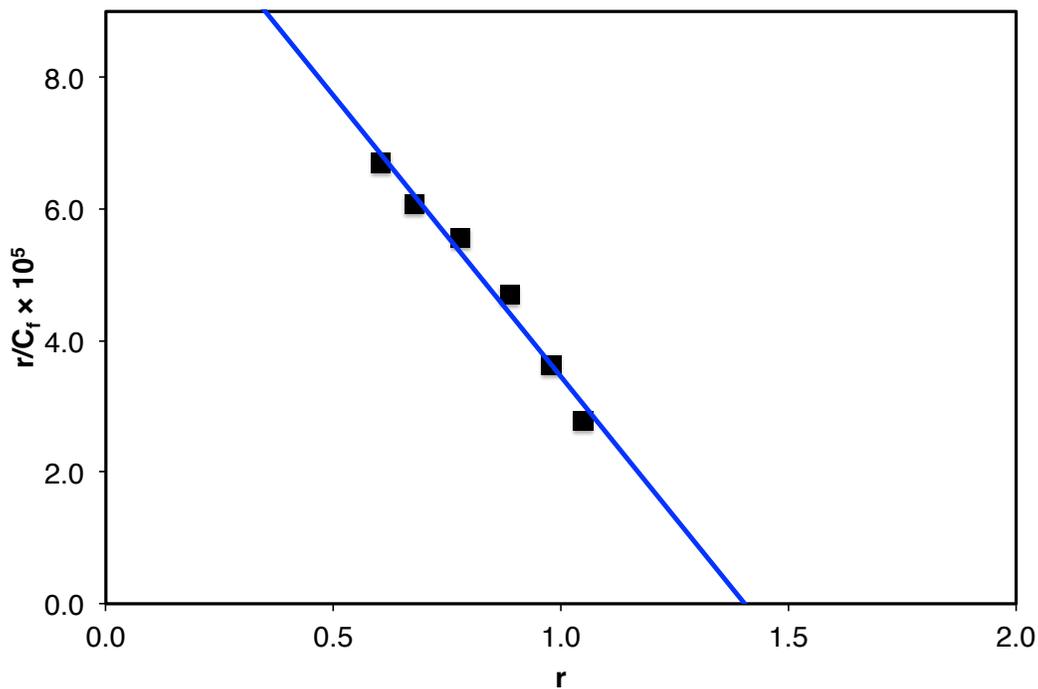
of absorbance or optical bleaching at 422 nm represents the removal of free ligand from solution. An optical bleaching plot demonstrates that ligand is almost fully bound at molar ratios greater than two for the artificial G-quadruplex binding ligand receptor **7** (Figure 5.12).

An end-stacking mode where one TMPyP4 molecule is π-π stacked onto the exposed face of the assembled G-tetrad of **7** is proposed (Figure 5.13).

Figure 5.11 Job plot for the binding of guanine conjugate **7** with TMPyP4 ligand in 10 mM

Tris-borate pH 8, 0.1 mM EDTA. The normalized change in absorbance at 436

nm is shown as a function of mol fraction of TMPyP4 for mixed solutions of **7** and

TMPyP4 with fixed concentration at 5 μM. 436 nm corresponds to the maximum

bathochromic shifted Soret band observed for bound TMPyP4.
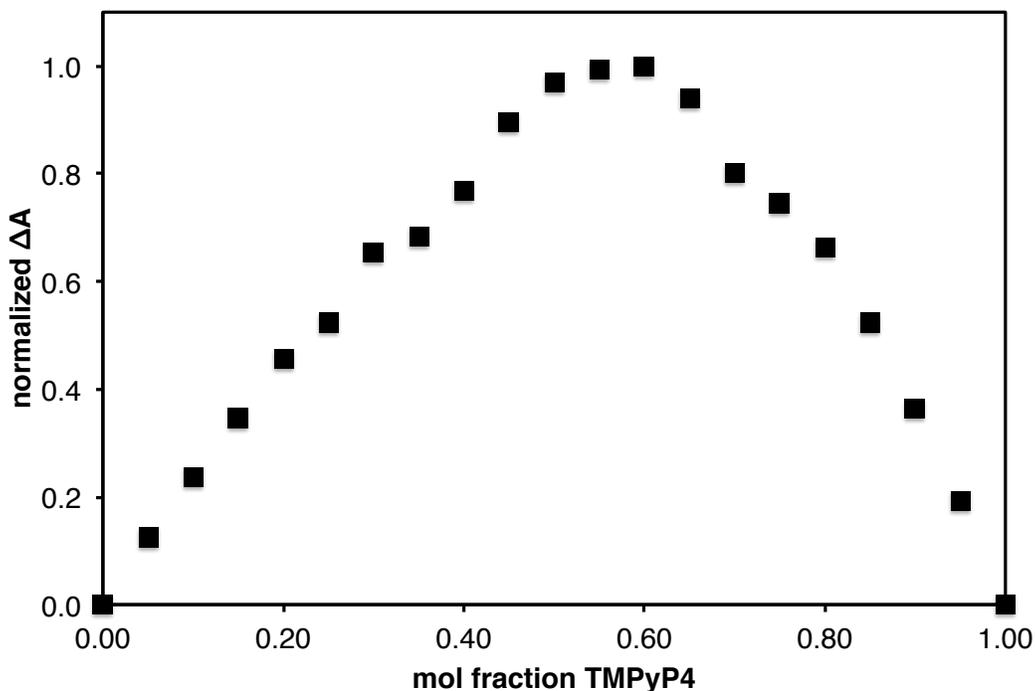
Some controversy exists in the literature with respect to the nature of binding of

TMPyP4 to G-quadruplex DNA with intercalative, end-stacking, and even loop-stacking

models of binding being proposed. Phan *et al.* provided an NMR solution model of a

terminally stacked TMPyP4 ligand with a parallel G-quadruplex formed from a c-myc

oncogene promoter region sequence.[72] In contrast, Wei and coworkers carried out

spectroscopic investigations that supported multiple intercalative sites on a self-folded

Figure 5.12   Bleaching plot for titration of 5 µM TMPyP4 ligand with guanine conjugate **7** in 10

mM Tris-borate pH 8, 0.1 mM EDTA monitored at 422 nm.
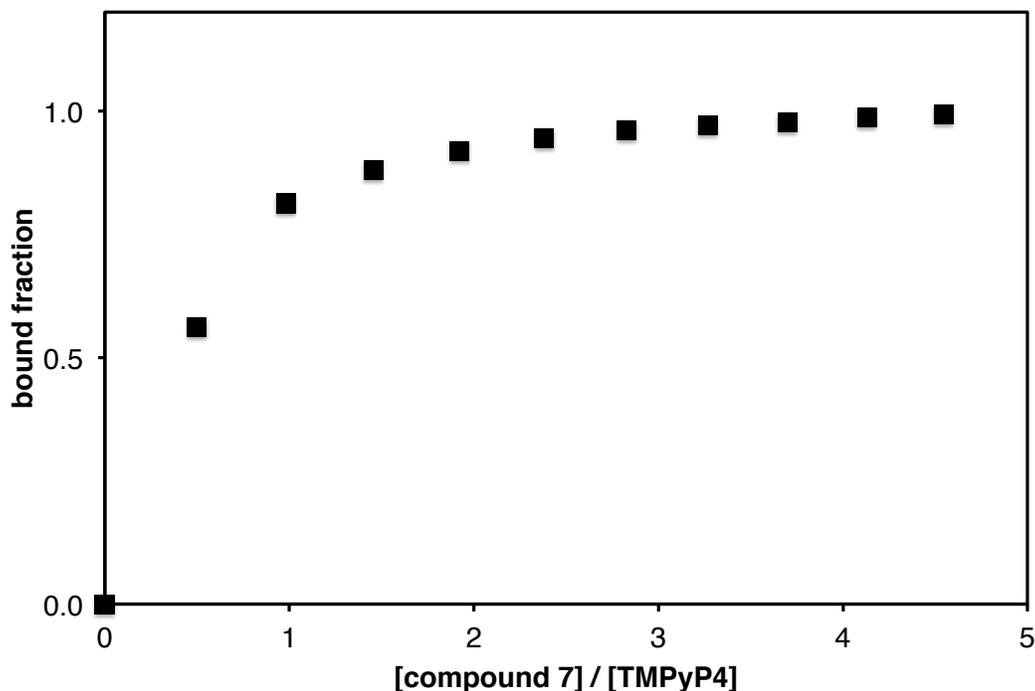
human telomere G-quadruplex.[73] In a crystal structure complex with a parallel dimer G-quadruplex containing a human telomere repeat, the TMPyP4 ligand was found to stack to a TTA loop without direct interaction with the guanine bases.[176]

It has been shown here that guanine based cavitand-nucleotide conjugate **7** is an effective receptor for several small molecule ligands with telomerase inhibition and G-quadruplex stabilization properties. Furthermore, **7** can be considered a model for the guanine tetrads in G-quadruplex DNA, for example of the human telomeric type, and may find practical applications for potential antitumor drug development in the telomerase targeting regime.

Figure 5.13   Cartoon representation of end-stacking of TMPyP4 ligand (orange) to the

exposed G-tetrad face (blue) of conjugate **7**. The cavitand (black) has its concave

surface directed toward the quartet.
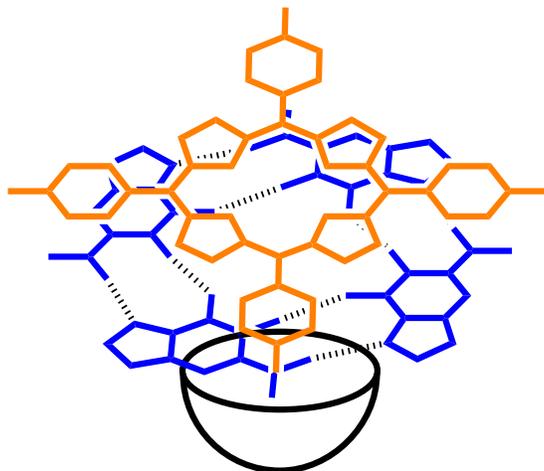
5.5   Experimental Procedures

Further information regarding CD, NMR, and ICP-MS experiments can be found in Section 4.6 Experimental Procedures. Procedures used in the construction of Scatchard, Job, and bleaching plots were adapted from those previously reported by Haq *et al.*[68] A 10 mM Tris-borate pH 8 and 0.1 mM EDTA buffer was used for all optical binding studies. Scatchard analysis was carried out by titrating stepwise 5 μL aliquots of stock solution of guanine **7** in 5 μM TMPyP4 into a 500 μL solution of 5 μM TMPyP4 in a 1 cm cell at 25 °C. The cell was incubated for 5 minutes after each addition. The fractional decrease in absorbance ΔA, at 422 nm was calculated from the relationship ΔA

= $(A_{free} - A) / (A_{free} - A_{sat})$, where A_{free} is the absorbance of the free porphyrin, A is the measured absorbance upon addition of **7**, and A_{sat} is the absorbance at saturation for 8:1 molar ratio of **7** to porphyrin. The fraction of bound porphyrin, α is taken directly as ΔA, and C_f is determined using $C_f = C(1 - α)$, where C is the fixed total porphyrin concentration at 5 μM. The binding ratio r, was calculated using r = $(C - C_f) / $ [**7**]. An equation of a straight line using the least-squares method was fit to the data points obtained from a simple Scatchard plot of r/C_f versus r.

Chapter 6: Thesis Summary and Conclusions

Foreseeing the *in vivo* importance of the guanine tetrad would have been difficult at the time Gellert proposed his historic arrangement of guanine tetramers in aqueous gels formed from solutions of guanylic acid and sodium chloride.[12] It is now understood that this characteristic cyclic arrangement of four hydrogen bonded guanine bases comprises the basic structural motif in G-quadruplex DNA, a putative biological DNA secondary structure unique from duplex B-DNA.[54] Guanine-rich regions in the human genome, including those localized to telomeres and gene promoters, have a potential to form G-quadruplex structures. Much attention has been given to the discovery of small molecule G-quadruplex binding ligands that can stabilize G-quadruplex DNA and in turn inhibit the human telomere maintenance enzyme, telomerase. Telomerase inhibition via G-quadruplex ligand binding represents a novel strategy towards antitumor drug design.

G-quadruplex DNA is intrinsically unstable and requires the presence of cations, typically sodium or potassium, for strand association. Furthermore, the guanine tetrad or quartet assembly itself is also unstable, and has never been reported to exist as an unstacked single entity in water. A synthetic question was asked as to whether a single isolated guanine quartet could exist in its native water environment, and this question formed the inspiration for the research presented in the preceding chapters. Template-assembly of nucleobases by application of a cavitand template molecule was utilized as a synthetic strategy.

In Chapter 2, classical DNA synthetic methodologies were adapted to accomplish the synthesis of water and methanol soluble phosphate-linked cavitand-nucleotide conjugates through phosphite triester and phosphoramidite couplings. All four DNA nucleotides were successfully coupled four-fold onto an individual cavitand template through a DNA-like 3' phosphate linkage. The first known synthesis of an individual substrate tetra-functionalized with guanylate was achieved. In addition, some of these compounds provided early examples of water-soluble tetra-guanine conjugates. The phosphoramidite approach simplified deblocking steps and typically gave improved yields. In contrast, the phosphite triester strategy eliminated the synthesis and isolation of unstable phosphoramidites that require expensive starting materials. These synthetic procedures may find application for multiple couplings of oligonucleotides and non-nucleic acid elements, such as of peptides and sugar molecules.

Chapter 3 described a synthetic route to water-soluble cavitand-guanosine conjugates through installation of triazole-linked nucleosides on cavitands with phosphates located at the pendant position. Evidence for guanine quartet template-assembly was not obvious in these compounds during preliminary characterization experiments. Design changes with respect to water-solubilizing group position and linker structure may improve nucleobase template-assembly in these compounds.

A rigorous characterization in methanol solvent for phosphate linked cavitand-nucleotide conjugates functionalized at the feet position was described in Chapter 4. Circular dichroism spectroscopy was demonstrated to be a useful technique for probing potential nucleobase association through template-assembly processes. Assembly of the guanine and thymine conjugates was interpreted to be more significant than of the

139

adenine and cytosine conjugates. Further characterization through NMR methods and ICP-MS gave support for the isolation of cation-free unimolecular template-assembled guanine and thymine quartets in methanol. The possibility of cation binding in these compounds remains to be explored. Further investigation into whether cavitand templates assemble rare adenine or novel cytosine quartets should be pursued.

In Chapter 5, characterization by CD, NMR, and ICP-MS experiments provided evidence that a water-soluble phosphate-linked cavitand-nucleotide conjugate functionalized at the rim position was template-assembling an isolated guanine quartet. This result constituted evidence for the first time of a cation-free unstacked guanine tetrad in water. The role of sodium or potassium ion binding to this construct was concluded to be ambiguous. This water-soluble template-assembled guanine quartet may serve as a minimal model for human telomeric DNA and is believed to possess an exposed G-quartet face suitable for docking G-quadruplex binding ligands. Optical binding studies demonstrated that this construct can in fact act as an artificial G-quadruplex binding ligand receptor. Binding to known small molecule G-quadruplex ligands and human telomerase inhibitors TMPyP4 and Piper was shown. Interactions of this guanine conjugate with other ligands should be explored. The synthesis of template-assembled G-quartet and G-quadruplex structures and the study of potential interactions between these structures with small molecule ligands may provide opportunities for developing anticancer therapeutics in the future.

References

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

(2) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964-967.

(3) Zamenhof, S.; Brawerman, G.; Chargaff, E. *Biochim. Biophys. Acta* **1952**, *9*, 402-405.

(4) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738-740.

(5) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740-741.

(6) Hoogsteen, K. *Acta Crystallogr.* **1959**, *12*, 822-823.

(7) Hoogsteen, K. *Acta Crystallogr.* **1963**, *16*, 907-916.

(8) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755-758.

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

(10) Bang, I. *Biochem. Z.* **1910**, *26*, 293.

(11) Ralph, R. K.; Khorana, H. G.; Connors, W. J. *J. Am. Chem. Soc.* **1962**, *84*, 2265-2266.

(12) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013-2018.

(13) Donohue, J. *Proc. Natl. Acad. Sci. U. S. A.* **1956**, *42*, 60-65.

(14) Zimmerman, S. B.; Cohen, G. H.; Davies, D. R. *J. Mol. Biol.* **1975**, *92*, 181-191.

(15) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I.; Blackburn, E. H. *Cell* **1987**, *51*, 899-908.

(16) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364-366.

(17) Williamson, J. R.; Rachuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871-880.

(18) Sundquist, W. I.; Klug, A. *Nature* **1989**, *342*, 825-829.

(19) Watson, J. D. *Nat. New Biol.* **1972**, *239*, 197-201.

(20) Olovniko.Am *J. Theor. Biol.* **1973**, *41*, 181-190.

(21) Levy, M. Z.; Allsopp, R. C.; Futcher, A. B.; Greider, C. W.; Harley, C. B. *J. Mol. Biol.* **1992**, *225*, 951-960.

(22) Sugino, A.; Hirose, S.; Okazaki, R. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 1863-1867.

(23) Blackburn, E. H.; Szostak, J. W. *Annu. Rev. Biochem.* **1984**, *53*, 163-194.

(24) Blackburn, E. H. *Nature* **1991**, *350*, 569-573.

(25) McClintock, B. *Genetics* **1941**, *26*, 234-282.

(26) Blackburn, E. H.; Gall, J. G. *J. Mol. Biol.* **1978**, *120*, 33-53.

(27) Szostak, J. W.; Blackburn, E. H. *Cell* **1982**, *29*, 245-255.

(28) Greider, C. W.; Blackburn, E. H. *Cell* **1985**, *43*, 405-413.

(29) Greider, C. W.; Blackburn, E. H. *Cell* **1987**, *51*, 887-898.

(30) Greider, C. W.; Blackburn, E. H. *Nature* **1989**, *337*, 331-337.

(31) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6622-6626.

(32) Morin, G. B. *Cell* **1989**, *59*, 521-529.

(33) Makarov, V. L.; Hirose, Y.; Langmore, J. P. *Cell* **1997**, *88*, 657-666.

(34) Feng, J. L., et al *Science* **1995**, *269*, 1236-1241.

(35) Palm, W.; de Lange, T. *Annu. Rev. Genet.* **2008**, *42*, 301-334.

(36) de Lange, T. *Genes Dev.* **2005**, *19*, 2100-2110.

(37) Kang, C.; Zhang, X. H.; Ratliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, *356*, 126-131.

(38) 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.

(39) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263-282.

(40) Parkinson, G.; Lee, M.; Neidle, S. *Nature* **2002**, *417*, 876-880.

(41) Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Pluckthun, A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8572-8577.

(42) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2007**, *35*, 406-413.

(43) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167-1172.

(44) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593-11598.

(45) Grand, C. L.; Han, H. Y.; Munoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Molecular Cancer Therapeutics* **2002**, *1*, 565-573.

(46) Hayflick, L.; Moorhead, P. S. *Exp. Cell Res.* **1961**, *25*, 585-620.

(47) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458-460.

(48) Counter, C. M.; Avilion, A. A.; Lefeuvre, C. E.; Stewart, N. G.; Greider, C. W.; Harley, C. B.; Bacchetti, S. *EMBO J.* **1992**, *11*, 1921-1929.

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

(50) Riou, J.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Megnin-Chanet, F.; Helene, C.; Mergny, J. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2672-2677.

(51) Norton, J. C.; Piatyszek, M. A.; Wright, W. E.; Shay, J. W.; Corey, D. R. *Nat. Biotechnol.* **1996**, *14*, 615-619.

(52) Sun, D.; Thompson, B.; Cathers, B.; Salazar, M.; Kerwin, S.; Trent, J.; Jenkins, T.; Neidle, S.; Hurley, L. *J. Med. Chem.* **1997**, *40*, 2113-2116.

(53) Franceschin, M. *Euro. J. Org. Chem.* **2009**, 2225-2238.

(54) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239-250.

(55) De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.; Mergny, J. *Biochimie* **2008**, *90*, 131-155.

(56) Monchaud, D.; Teulade-Fichou, M. *Org. Biomol. Chem.* **2008**, *6*, 627-636.

(57) Ou, T.; Lu, Y.; Tan, J.; Huang, Z.; Wong, K.; Gu, L. *ChemMedChem* **2008**, *3*, 690-713.

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

(59) Neidle, S.; Parkinson, G. *Nat. Rev. Drug Discovery* **2002**, *1*, 383-393.

(60) Han, H. Y.; Hurley, L. H. *Trends Pharmacol. Sci.* **2000**, *21*, 136-142.

(61) Neidle, S.; Read, M. A. *Biopolymers* **2000**, *56*, 195-208.

(62) Harrison, R.; Gowan, S.; Kelland, L.; Neidle, S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2463-2468.

(63) Gowan, S.; Harrison, J.; Patterson, L.; Valenti, M.; Read, M.; Neidle, S.; Kelland, L. *Mol. Pharmacol.* **2002**, *61*, 1154-1162.

(64) Haider, S.; Parkinson, G.; Neidle, S. *J. Mol. Biol.* **2003**, *326*, 117-125.

(65) Fedoroff, O.; Salazar, M.; Han, H.; Chemeris, V.; Kerwin, S.; Hurley, L. *Biochemistry* **1998**, *37*, 12367-12374.

(66) Anantha, N.; Azam, M.; Sheardy, R. *Biochemistry* **1998**, *37*, 2709-2714.

(67) Wheelhouse, R.; Sun, D.; Han, H.; Han, F.; Hurley, L. *J. Am. Chem. Soc.* **1998**, *120*, 3261-3262.

(68) Haq, I.; Trent, J.; Chowdhry, B.; Jenkins, T. *J. Am. Chem. Soc.* **1999**, *121*, 1768-1779.

(69) Dixon, I. M.; Lopez, F.; Tejera, A. M.; Esteve, J.; Blasco, M. A.; Pratviel, G.; Meunier, B. *J. Am. Chem. Soc.* **2007**, *129*, 1502-1503.

(70) Han, H.; Langley, D.; Rangan, A.; Hurley, L. *J. Am. Chem. Soc.* **2001**, *123*, 8902-8913.

(71) Goncalves, D. P. N.; Rodriguez, R.; Balasubramanian, S.; Sanders, J. K. M. *Chem. Commun.* **2006**, 4685-4687.

(72) Phan, A.; Kuryavyi, V.; Gaw, H.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *1*, 167-173.

(73) Wei, C.; Jia, G.; Yuan, J.; Feng, Z.; Li, C. *Biochemistry* **2006**, *45*, 6681-6691.

(74) Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. *J. Am. Chem. Soc.* **2001**, *123*, 1262-1263.

(75) Kim, M.; Vankayalapati, H.; Kazuo, S.; Wierzba, K.; Hurley, L. *J. Am. Chem. Soc.* **2002**, *124*, 2098-2099.

(76) Tauchi, T.; Shin-ya, K.; Sashida, G.; Sumi, M.; Nakajima, A.; Shimamoto, T.; Ohyashiki, J. H.; Ohyashiki, K. *Oncogene* **2003**, *22*, 5338-5347.

(77) Doi, T.; Yoshida, M.; Shin-ya, K.; Takahashi, T. *Org. Lett.* **2006**, *8*, 4165-4167.

(78) Naasani, I.; Seimiya, H.; Yamori, T.; Tsuruo, T. *Cancer Res.* **1999**, *59*, 4004-4011.

(79) Franceschin, M.; Rossetti, L.; D'Ambrosio, A.; Schirripa, S.; Bianco, A.; Ortaggi, G.; Savino, M.; Schultes, C.; Neidle, S. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1707-1711.

(80) Gomez, D.; Mergny, J. L.; Riou, J. F. *Cancer Res.* **2002**, *62*, 3365-3368.

(81) De Cian, A.; Cristofari, G.; Reichenbach, P.; De Lemos, E.; Monchaud, D.; Teulade-Fichou, M.; Shin-Ya, K.; Lacroix, L.; Lingner, J.; Mergny, J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 17347-17352.

(82) Nagai, K.; Hayakawa, K.; Ukai, S.; Kanematsu, K. *J. Org. Chem.* **1986**, *51*, 3931-3939.

(83) Kim, M. S.; Gokel, G. W. *Chem. Commun.* **1987**,1686-1688.

(84) Schall, O. F.; Gokel, G. W. *J. Am. Chem. Soc.* **1994**, *116*, 6089-6100.

(85) Sessler, J. L.; Wang, R. Z. *J. Am. Chem. Soc.* **1996**, *118*, 9808-9809.

(86) Sessler, J. L.; Wang, R. Z. *J. Org. Chem.* **1998**, *63*, 4079-4091.

(87) Mutter, M.; Vuilleumier, S. *Angew. Chem. Int. Ed.* **1989**, *28*, 535-554.

(88) Consoli, G. M. L.; Granata, G.; Garozzo, D.; Mecca, T.; Geraci, C. *Tetrahedron Lett.* **2007**, *48*, 7974-7977.

(89) 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.

(90) Sidorov, V.; Kotch, F. W.; El-Khouedi, M.; Davis, J. T. *Chem. Commun.* **2000**, 2369-2370.

(91) Nikan, M.; Sherman, J. C. *Angew. Chem. Int. Ed.* **2008**, *47*, 4900-4902.

(92) Nikan, M.; Sherman, J. C. *J. Org. Chem.* **2009**, *74*, 5211-5218.

(93) Oliviero, G.; Amato, J.; Borbone, N.; Galeone, A.; Petraccone, L.; Varra, M.; Piccialli, G.; Mayol, L. *Bioconjug. Chem.* **2006**, *17*, 889-898.

(94) Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. *Tetrahedron Lett.* **2004**, *45*, 4869-4872.

(95) Oliviero, G.; Borbone, N.; Amato, J.; D'Errico, S.; Galeone, A.; Piccialli, G.; Varra, M.; Mayol, L. *Biopolymers* **2009**, *91*, 466-477.

(96) Oliviero, G.; Amato, J.; Borbone, N.; D'Errico, S.; Galeone, A.; Mayol, L.; Haider, S.; Olubiyi, O.; Hoorelbeke, B.; Balzarini, J.; Piccialli, G. *Chem. Commun.* **2010**, *46*, 8971-8973.

(97) Murat, P.; Cressend, D.; Spinelli, N.; Van der Heyden, A.; Labbe, P.; Dumy, P.; Defrancq, E. *ChemBioChem* **2008**, *9*, 2588-2591.

(98) Murat, P.; Bonnet, R.; Van der Heyden, A.; Spinelli, N.; Labbe, P.; Monchaud, D.; Teulade-Fichou, M.; Dumy, P.; Defrancq, E. *Chem. Eur. J.* **2010**, *16*, 6106-6114.

(99) Cheong, C. J.; Moore, P. B. *Biochemistry* **1992**, *31*, 8406-8414.

(100) Deng, J.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13665-13670.

(101) Witkowski, H.; Freisinger, E.; Lippert, B. *Chem. Commun.* **1997**,1315-1316.

(102) Xu, Y.; Ishizuka, T.; Kimura, T.; Komiyama, M. *J. Am. Chem. Soc.* **2010**, *132*, 7231-7233.

(103) Hui, B. W.; Sherman, J. C. *Chem. Commun.* **2012**, *48*, 109-111.

(104) Sket, P.; Plavec, J. *J. Am. Chem. Soc.* **2010**, *132*, 12724-12732.

(105) Searle, M.; Williams, H.; Gallagher, C.; Grant, R.; Stevens, M. *Org. Biomol. Chem.* **2004**, *2*, 810-812.

(106) Patel, P. K.; Koti, A. S. R.; Hosur, R. V. *Nucleic Acids Res.* **1999**, *27*, 3836-3843.

(107) Michelson, A. M.; Todd, A. R. *J. Chem. Soc.* **1955**, 2632-2638.

(108) Agarwal, K. L.; Buchi, H.; Caruthers, M. H.; Gupta, N.; Khorana, H. G.; Kleppe, K.; Kumar, A.; Ohtsuka, E.; Rajbhand U. L.; Vandesan J.H.; Sgaramel, V; Weber, H.; Yamada, T. *Nature* **1970**, *227*, 27-34.

(109) Letsinger, R. L.; Lunsford, W. B. *J. Am. Chem. Soc.* **1976**, *98*, 3655-3661.

(110) Beaucage, S.; Caruthers, M. *Tetrahedron Lett.* **1981**, *22*, 1859-1862.

(111) Caruthers, M. H. *Science* **1985**, *230*, 281-285.

(112) Ogilvie, K. K. *Can. J. Chem.* **1973**, *51*, 3799-3807.

(113) Ti, G. S.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 1316-1319.

(114) Eisenhuth, R.; Richert, C. *J. Org. Chem.* **2009**, *74*, 26-37.

(115) Mezo, A. R.; Sherman, J. C. *J. Org. Chem.* **1998**, *63*, 6824-6829.

(116) Khorana, H. G. *Pure Appl. Chem.* **1968**, *17*, 349.

(117) Ogilvie, K. K.; Theriault, N. Y.; Seifert, J. M.; Pon, R. T.; Nemer, M. J. *Can J. Chem.* **1980**, *58*, 2686-2693.

(118) Adamiak, R. W.; Biala, E.; Grzeskowiak, K.; Kierzek, R.; Kraszewski, A.; Markiewicz, W. T.; Stawinski, J.; Wiewiorowski, M. *Nucleic Acids Res.* **1977**, *4*, 2321-2329.

(119) Fraser, J. R.; Borecka, B.; Trotter, J.; Sherman, J. C. *J. Org. Chem.* **1995**, *60*, 1207-1213.

(120) Arnaudneu, F.; Collins, E. M.; Deasy, M.; Ferguson, G.; Harris, S. J.; Kaitner, B.; Lough, A. J.; Mckervey, M. A.; Marques, E.; Ruhl, B. L.; Schwingweill, M. J.; Seward, E. M. *J. Am. Chem. Soc.* **1989**, *111*, 8681-8691.

(121) Weber, H.; Khorana, H. G. *J. Mol. Biol.* **1972**, *72*, 219-249.

(122) Sinha, N.; Biernat, J.; Koster, H. *Tetrahedron Lett.* **1983**, *24*, 5843-5846.

(123) Nikan, M.; Bare, G. A. L.; Sherman, J. C. *Tetrahedron Lett.* **2011**, *52*, 1791-1793.

(124) Biros, S. M.; Rebek, J., Jr. *Chem. Soc. Rev.* **2007**, *36*, 93-104.

(125) Nikan, M. Dissertation, University of British Columbia, 2009.

(126) Gui, X.; Sherman, J. *Chem. Commun.* **2001**, 2680-2681.

(127) Gibb, B.; Chapman, R.; Sherman, J. *J. Org. Chem.* **1996**, *61*, 1505-1509.

(128) Sherman, J.; Knobler, C.; Cram, D. *J. Am. Chem. Soc.* **1991**, *113*, 2194-2204.

(129) Cram, D.; Jaeger, R.; Deshayes, K. *J. Am. Chem. Soc.* **1993**, *115*, 10111-10116.

(130) Stout, M.; Robins, M.; Olsen, R.; Robins, R. *J. Med. Chem.* **1969**, *12*, 658-662.

(131) Liu, F.; Austin, D. *J. Org. Chem.* **2001**, *66*, 8643-8645.

(132) Dean, D. *Synth. Commun.* **2002**, *32*, 1517-1521.

(133) Park, T.; Todd, E.; Nakashima, S.; Zimmerman, S. *J. Am. Chem. Soc.* **2005**, *127*, 18133-18142.

(134) Kolb, H.; Finn, M.; Sharpless, K. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.

(135) Amblard, F.; Cho, J. H.; Schinazi, R. F. *Chem. Rev.* **2009**, *109*, 4207-4220.

(136) Denny, W.; Leupin, W.; Kearns, D. *Helv. Chim. Acta* **1982**, *65*, 2372-2393.

(137) Sreenivasachary, N.; Lehn, J. *Chem. Asian J.* **2008**, *3*, 134-139.

(138) Sreenivasachary, N.; Lehn, J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5938-5943.

(139) Chantot, J.; Sarocchi, M.; Guschlbauer.W *Biochimie* **1971**, *53*, 347-354.

(140) Murat, P.; Gennaro, B.; Garcia, J.; Spinelli, N.; Dumy, P.; Defrancq, E. *Chem. Eur. J.* **2011**, *17*, 5791-5795.

(141) Gansey, M.; Bakker, F.; Feiters, M.; Geurts, H.; Verboom, W.; Reinhoudt, D. *Tetrahedron Lett.* **1998**, *39*, 5447-5450.

(142) Masiero, S.; Trotta, R.; Pieraccini, S.; De Tito, S.; Perone, R.; Randazzo, A.; Spada, G. P. *Org. Biomol. Chem.* **2010**, *8*, 2683-2692.

(143) Nikan, M.; Sherman, J. C. unpublished results, 2009.

(144) Berova, N.; Di Bari, L.; Pescitelli, G. *Chem. Soc. Rev.* **2007**, *36*, 914-931.

(145) Nakanishi, K.; Berova, N.; Woody, R. W., Eds.; In *Circular Dichroism: Principles and Applications*; Wiley-VCH: New York, 2000; Vol. 2nd ed.

(146) Fasman, G. D., Ed.; In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press: New York, 1996.

(147) Paramasivan, S.; Rujan, I.; Bolton, P. H. *Methods* **2007**, *43*, 324-331.

(148) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.

(149) Dapic, V.; Abdomerovic, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J.; Bates, P. *Nucleic Acids Res.* **2003**, *31*, 2097-2107.

(150) Voelter, W.; Records, R.; Bunnenbe.E; Djerassi, C. *J. Am. Chem. Soc.* **1968**, *90*, 6163-6170.

(151) Jin, R.; Gaffney, B.; Wang, C.; Jones, R.; Breslauer, K. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 8832-8836.

(152) Scaria, P.; Shire, S.; Shafer, R. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 10336-10340.

(153) Feigon, J.; Denny, W.; Leupin, W.; Kearns, D. *Biochemistry* **1983**, *22*, 5930-5942.

(154) Gueron, M.; Kochoyan, M.; Leroy, J. *Nature* **1987**, *328*, 89-92.

(155) Kochoyan, M.; Leroy, J.; Gueron, M. *J. Mol. Biol.* **1987**, *196,* 599-609.

(156) Kessler, H.; Gehrke, M.; Griesinger, C. *Angew. Chem. Int. Ed.* **1988**, *27*, 490-536.

(157) Bothnerby, A.; Stephens, R.; Lee, J.; Warren, C.; Jeanloz, R. *J. Am. Chem. Soc.* **1984**, *106*, 811-813.

(158) Sket, P.; Plavec, J. *J. Am. Chem. Soc.* **2010**, *132*, 12724-12732.

(159) Cohen, Y.; Avram, L.; Frish, L. *Angew. Chem. Int Ed.* **2005**, *44*, 520-554.

(160) Johnson, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203-256.

(161) Kaucher, M.; Lam, Y.; Pieraccini, S.; Gottarelli, G.; Davis, J. *Chem. Euro. J.* **2005**, *11*, 164-173.

(162) Wong, A.; Ida, R.; Spindler, L.; Wu, G. *J. Am. Chem. Soc.* **2005**, *127*, 6990-6998.

(163) Kaucher, M.; Lam, Y.; Pieraccini, S.; Gottarelli, G.; Davis, J. *Chem. Euro. J.* **2005**, *11*, 164-173.

(164) Stejskal, E.; Tanner, J. *J. Chem. Phys.* **1965**, *42*, 288-292.

(165) Waldeck, A.; Kuchel, P.; Lennon, A.; Chapman, B. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *30*, 39-68.

(166) Davis, J. T. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.

(167) Otero, R.; Schock, M.; Molina, L.; Laegsgaard, E.; Stensgaard, I.; Hammer, B.; Besenbacher, F. Angew. *Angew. Chem. Int. Ed.* **2005**, *44*, 2270-2275.

(168) Sessler, J.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. *Angew. Chem. Int. Ed.* 2000, 39, 1300-1303.

(169) Kettani, A.; Gueron, M.; Leroy, J. *J. Am. Chem. Soc.* **1997**, *119*, 1108-1115.

(170) Boelens, R.; Scheek, R. M.; Dijkstra, K.; Kaptein, R. *J. Magn. Reson.* **1985**, *62,* 378-386.

(171) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.

(172) Yue, D. J. E.; Lim, K. W.; Anh Tuan Phan *J. Am. Chem. Soc.* **2011**, *133*, 11462-11465.

(173) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723-2735.

(174) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963-9970.

(175) Qin, Y.; Hurley, L. H. *Biochimie* **2008**, *90*, 1149-1171.

(176) Parkinson, G. N.; Ghosh, R.; Neidle, S. *Biochemistry* **2007**, *46*, 2390-2397.

(177) Han, F.; Wheelhouse, R.; Hurley, L. *J. Am. Chem. Soc.* **1999**, *121*, 3561-3570.

(178) Mazzitelli, C.; Brodbelt, J.; Kern, J.; Rodriguez, M.; Kerwin, S. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593-604.

(179) Rossetti, L.; Franceschin, M.; Bianco, A.; Ortaggi, G.; Savino, M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2527-2533.

(180) Kerwin, S. M.; Chen, G.; Kern, J. T.; Thomas, P. W. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 447-450.

(181) Han, H.; Cliff, C.; Hurley, L. *Biochemistry* **1999**, *38*, 6981-6986.

(182) Jenkins, T. C. *Methods Mol. Biol.* **1997**, *90*, 195-218.