CONSTRUCTION OF TEMPLATE-ASSEMBLED PYRIMIDINE-BASED QUARTETS AND QUADRUPLEXES

by

Benjamin Wei Qiang Hui

B.Sc., National University of Singapore, 2007

M.Sc., Nanyang Technological University, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2014

Abstract

Pyrimidine-based quartets and quadruplexes are unstable and thus are rarely encountered in nature. Uracil (U) and thymine (T) quartets in the solution state have only been found as part of pre-existing G-quadruplex scaffolds and the corresponding quadruplexes have not been reported. Studies on such systems might shed light on their role in nucleic acid topology and stability. This thesis describes the assembly and structural characterization of these motifs *in vitro* as a result of grafting the respective nucleosides onto resorcinol-based cavitands. These rigid macrocycles serve as molecular templates on which these motifs are preorganized. Reduction of entropic loss improves thermodynamic stability and promotes self-assembly.

A convergent synthetic strategy was employed for accessing these cavitand-nucleoside conjugates. Cavitands and nucleosides were prepared separately using established literature methods, and the final coupling step of the two components entailed a copper (I)-catalyzed azide-alkyne cycloaddition, or a "click" reaction. NMR spectroscopy was used extensively in signal assignment, structure elucidation and oligomeric state analysis. CD spectroscopy was employed in some cases to provide further confirmation of defined structure.

Findings indicated the spontaneous self-assembly of a U-quartet in $CDCl_3$ at both 25 ºC and –20 ºC. In the presence of a metal cation (Sr^{2+}), symmetric homodimerization of two U-quartets occurs at 25 ºC. The corresponding U-quadruplex unit was identified in DMSO-d_6 at 25 ºC. The T-quartet was shown to be nonexistent at 25 ºC, but assembles at a low temperature of –40 ºC.

No evidence for metal cation uptake was found at 25 ºC. Assembly of the T-quadruplex was confirmed in DMSO-d_6 at 25 ºC. In all of these systems, stacking of the nucleobase and triazole linker rings was indicated suggesting π-stacking interactions to be a significant contributor to overall stability.

Preface

The directions and goals of this thesis were discussed and agreed upon between the author and research supervisor. All work described in this thesis was performed solely by the author. A version of Chapter 2 has been published. Benjamin Wei-Qiang Hui and John C. Sherman (2012) Synthesis and characterization of a template-assembled synthetic U-quartet. Chem. Comm. 48:109-111. The first draft of the manuscript and all work described therein was completed by the author. A version of Chapter 3 has been published. Benjamin Wei-Qiang Hui and John C. Sherman (2012) A template-assembled synthetic U-quadruplex. ChemBioChem 13:1865-1868. The first draft of the manuscript and all work described therein was completed by the author. A version of Chapter 4 has been published. Benjamin Wei-Qiang Hui and John C. Sherman (2013) Self-assembly of a thymine quartet and quadruplex via an organic template. Tetrahedron Letters 55:1479-1485. The first draft of the manuscript and all work therein was completed by the author.

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
br	broad
Bu	butyl
C	cytosine
CD	circular dichroism
COSY	correlation nuclear magnetic resonance spectroscopy
d	doublet
dd	doublet of doublets
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMSO-d_6	deuterated dimethylsulfoxide
DNA	deoxyribonucleic acid
DOSY	diffusion-ordered nuclear magnetic resonance spectroscopy
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
G	guanine
GMP	guanosine monophosphate

h	hour
HRMS	high resolution mass spectrometry
m	multiplet
Me	methyl
MALDI-TOF	matrix-assisted laser desorption-ionization, time-of-flight
MS	mass spectrometry
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect nuclear magnetic resonance spectroscopy
NMR	nuclear magnetic resonance spectroscopy
PDB	Protein Databank
RNA	ribonucleic acid
s	singlet
rt	room temperature
t	triplet
td	triplet of doublets
T	thymine
TASP	template-assembled synthetic protein
TASQ	template-assembled synthetic G-quartet
TBAF	tetra-*n*-butylammonium fluoride
TBS	*tert*-butyldimethylsilyl
THF	tetrahydrofuran
U	uracil
VT	variable-temperature nuclear magnetic resonance spectroscopy

Acknowledgements

I would like to extend my appreciation to my advisor, Professor John Sherman for helpful discussions and providing me the opportunity to work in his laboratory. I also thank UBC and the Department of Chemistry for financial support in the form of scholarships. My fellow labmates Grant Bare, Jon Freeman and Hui Yang provided much needed support and inspiration during this time, and I am most fortunate to have them as friends. Special thanks go to the UBC NMR and mass spectrometry lab staff, namely Maria Ezhova, Zorana Danilovic, Yun Ling and Marshall Lapawa for their expert assistance.

Most of all, I am grateful to my parents for their unrelenting spiritual, moral and financial support throughout all these years.

Dedication

To my parents, and the Creator.

Chapter 1: Introduction

1.1 Beyond the covalent bond

Supramolecular chemistry is the study of non-covalent interactions within macromolecular systems and how these forces dictate structure.[1, 2] Specifically, the understanding of how discrete molecular subunits associate through these forces into cohesive and functional wholes is sought.[3, 4]

In the pursuit of understanding Nature at the molecular level, the *in vitro* construction of model biomolecule analogs for study is a persistent challenge. Template-directed self-assembly has emerged as a useful technique allowing facile access to such inherently complex molecular architectures.[5] This is achieved by the use of template molecules on which self-assembling components are attached, bringing them into close proximity for interaction, thus alleviating the entropic penalty and consequently increasing the thermodynamic stability of the assembled motifs. Of the numerous types of template molecules designed, from solid state substrates like carbon nanotubes[6, 7] and polymer nanofibers[8] to biomolecules like peptides[9] and DNA[10], the utility of organic resorcinol-based cavitands is noteworthy. These compounds represent a broader spectrum of macrocycles with an innate central cavity that are easily synthesized and derivatized, offer excellent chemical stability, and are structurally rigid.[11] In addition, the presence of at least a four-fold symmetry allows varying degrees of molecular complexity to be constructed on their surfaces. Our group has made contributions to the field of *de novo* protein design by the preparation of template-assembled synthetic proteins (TASPs) composed of medium-length peptides covalently bound to cavitand templates. Studies have revealed elevated thermal and kinetic stability of these TASPs vis-à-vis other synthetic systems.[12, 13] Slight modification of the

amino acid sequences was found to impart hydrolytic activity to the TASPs towards small molecule substrates.[14] This has possible implications for their use as enzyme mimics and is highly advantageous due to an expanded substrate scope and negligible loss of activity over time.

1.2 Introduction to DNA structure

1.2.1 Historical background

The successful isolation of DNA in the late 19th century was met with little enthusiasm due to its hitherto unknown function. By the early 20th century, the canonical nitrogenous bases had been identified (Figure 1.1a). In 1919, Levene characterized the nitrogenous base, 2'-deoxyribose sugar ring and the phosphate moieties and correctly recognized that these constitute a DNA nucleotide (Figure 1.1b).[15] The relative positions of the substituents on the sugar ring were as yet unclear. Extensive as these studies were at the time, there was still no reason to believe that DNA played any part in the transmission of heredity, that role being attributed to proteins as they were thought to be more chemically complex. This mindset persisted until the middle of the century when several groundbreaking experiments proved otherwise. In 1943, Avery, MacLeod and McCarthy, working with the *S. pneumoniae* bacterium, found that a mixture of live non-virulent Type II-R and killed virulent Type III-S bacterial particles was infective.[16] They deduced the only way this could happen was a transfer of genetic information from the virulent to the non-virulent bacteria, effecting their transformation to the virulent type. The agent responsible for this process was isolated and found to be DNA. Definitive evidence that DNA is indeed the hereditary material was uncovered in 1952, when Hershey and Chase found that bacteriophages transfer only their DNA to the host cells at the moment of infection.[17]

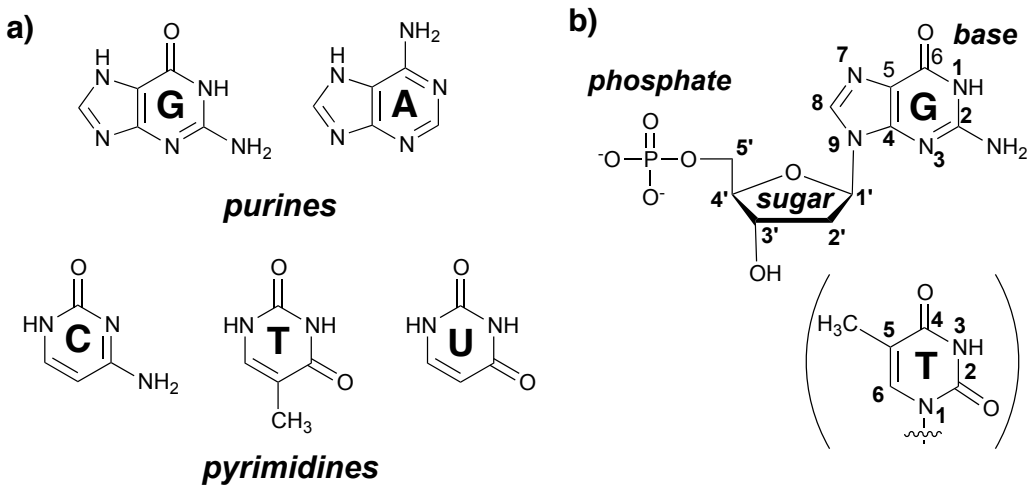


Figure 1.1 a) The canonical nucleobases (uracil replaces thymine in RNA). b) DNA nucleotide units consisting of the base, 2'-deoxyribose sugar and 5'-phosphate groups.

The structure of DNA was finally reported in 1953 by Watson and Crick through X-ray diffraction studies on DNA fibers.[18, 19] In this monumental work, they elucidated the double helical structure of DNA, comprising two antiparallel polynucleotide strands winding in a right-handed fashion about a lengthwise axis (Figure 1.2). In accordance with Chargaff's rule necessitating a universal 1:1 ratio of purine to pyrimidine nitrogenous bases in DNA,[20] they formulated the Watson-Crick complementary base pairing system where adenine (A) pairs with thymine (T), and guanine (G) with cytosine (C) via hydrogen bonds. Within their model, now recognized as B-DNA and the physiological form, the plane of each base pair is perpendicular to the helical axis. The helix twists 36º per base pair, and thus completes a turn every ten base pairs. A rise of 34 Å per turn is consistent with an individual base pair thickness of 3.4 Å. These geometric features confer a topological repeat pattern of major and minor grooves throughout the length of the molecule of fundamental importance to its function.

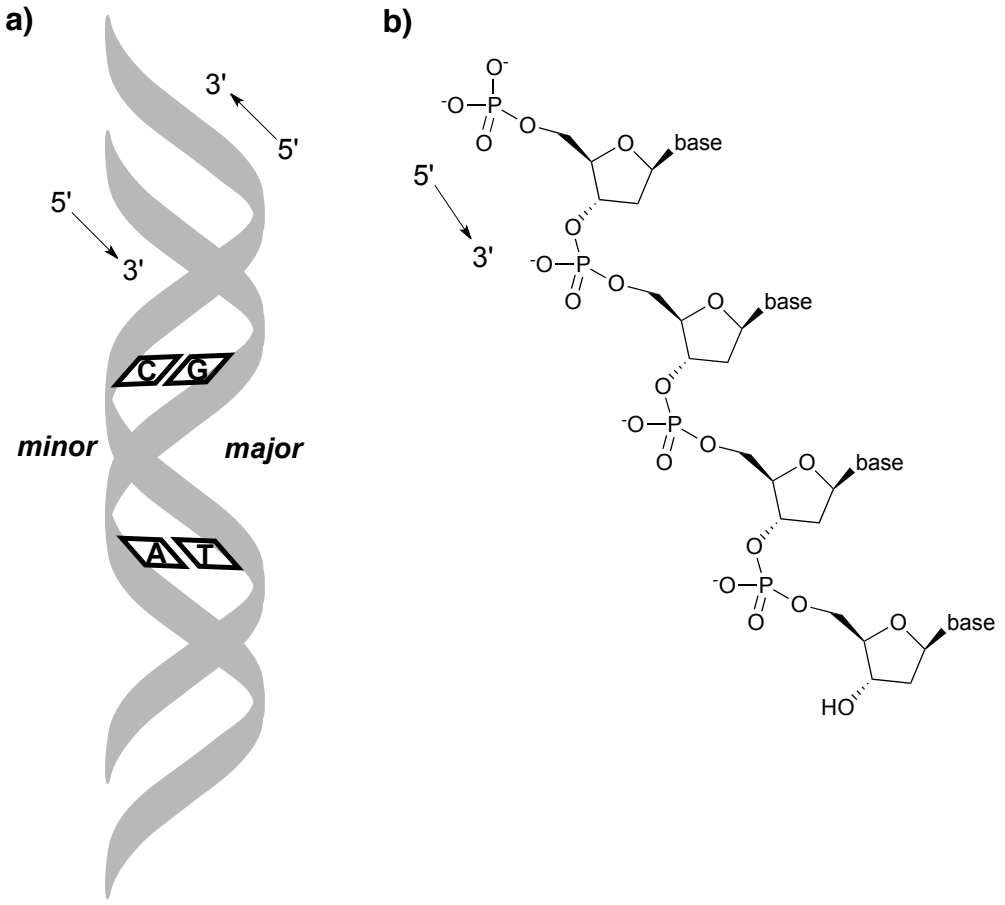


Figure 1.2 a) The double helix of B-DNA proposed by Watson and Crick. A•T and C•G base pairs are depicted along with the minor and major grooves and directional sense of both strands. b) Molecular structure of a generic oligonucleotide.

The implications of these findings were profound. Firstly, Watson-Crick base pairing all but precluded the possibility of a static, unchanging base sequence as assumed previously. The pretext for the origin of diversity of all terrestrial life was hence laid. Secondly, the closed double helix renders the bases unexposed to the environment. Therefore, a mechanism for DNA replication and expression yet unknown had to be elucidated. Inadvertently, the floodgates of

inquiry that were opened with this monumental discovery led to the modern fields of molecular biology and genetics.

1.2.2 Overview of nucleotide structure

In a typical 2'-deoxyribonucleotide, the nitrogenous base is bound to the sugar via a β-*N*-glycosidic bond connecting the anomeric C1' of the furanose ring and the N1 or N9 position of the pyrimidine or purine base respectively. Rotation about this bond allows for two possible spatial conformations: the *anti*- and *syn*- forms (Figure 1.3). In native DNA, the *anti* conformer is adopted because the *syn* conformer extends the base over the surface of the sugar ring resulting in steric interactions.

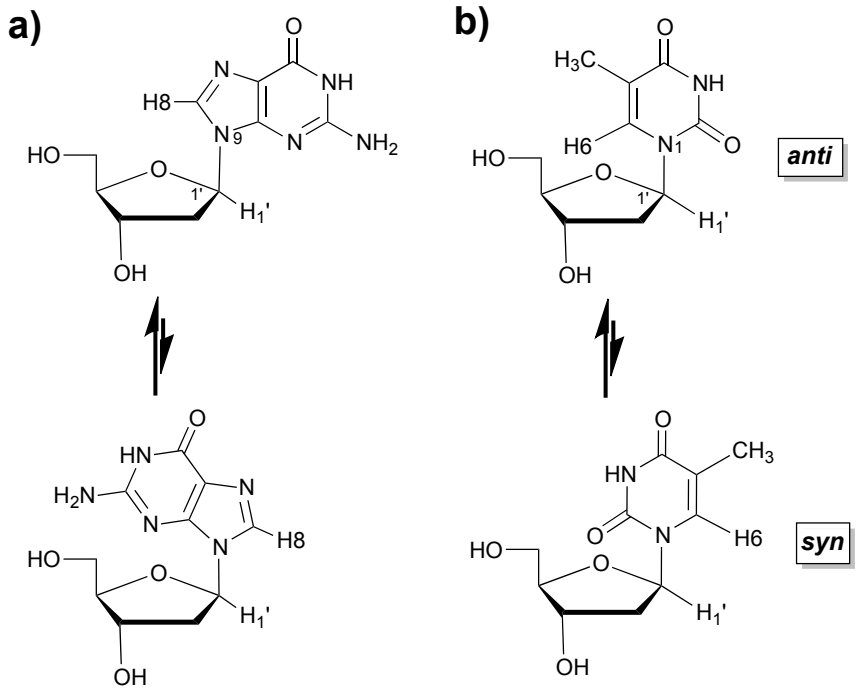


Figure 1.3 The *syn* and *anti* forms of a) adenosine and b) thymidine nucleosides.

Puckering of the ribose sugar ring to assume an envelope-like conformation causes either the C2' or C3' atom to extend out of the ring plane on the same face as C5' (Figure 1.4). In the C2'-*endo* conformation, the resultant relative positions of the 5'- and 3'-phosphate linkers gives rise to B-DNA. The C3'-*endo* species provides the A-DNA form, not normally found under physiological conditions.

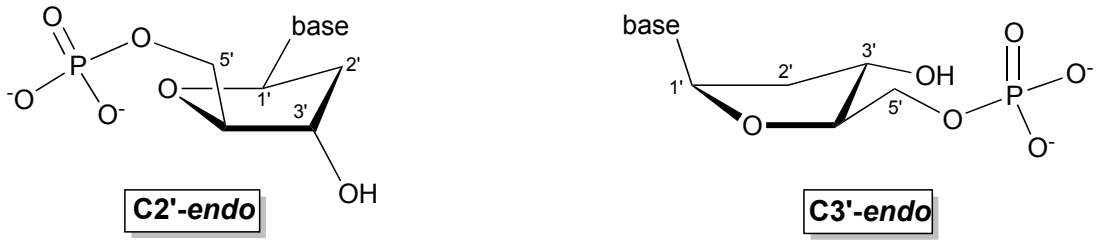


Figure 1.4 Ribose sugar ring puckering at either C2' or C3' positions.

1.2.3 Natural and unnatural base pairing in DNA

The Watson-Crick base pairing system helped codify organization within the large DNA molecule. A•T base pairs consist of two hydrogen bonds connecting the N1 and amino proton of adenine to H3 and O4 of thymine respectively. The G•C base pair entails three hydrogen bonds connecting the H1, amino proton and O6 of guanine to N3, O2 and amino proton of cytosine respectively. (Figure 1.5)

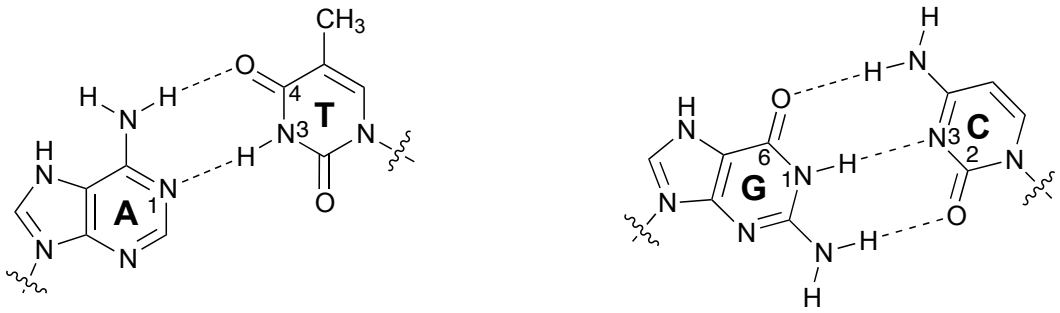


Figure 1.5 Canonical base pairs of DNA.

Tautomerization is not pronounced in the bases under physiological conditions although the presence of transferable protons allows for it. Guanine and thymine exist predominantly in the keto form, while adenine and cytosine assume the amino form. Gas phase studies conducted on uracil and thymine at elevated temperatures found only trace amounts of the enol tautomer.[21] This is noteworthy considering the aromaticity of the heterocyclic ring in the enol. Nonetheless, enolization results in a change of molecular conformation as well as the configuration of the hydrogen bond contacts. As a result, enolic thymine forms an anomalous base pair with guanine possessing three hydrogen bonds, resulting in a point mutation in DNA. Likewise, the minor imino tautomer of cytosine forms two hydrogen bonds with adenine instead of the usual three with guanine. (Figure 1.6a)

Interestingly, standard Watson-Crick hydrogen bonding modes do not always apply to canonical base pairs. Crystalline structures of A•T in the form of free bases 9-methyladenine and 1-methylthymine reveal the participation of the adenine N7 instead of N1 (Figure 1.6b).[22, 23] This pattern is known as Hoogsteen hydrogen bonding and is especially involved in the stabilization of guanine-based supramolecular assemblies to be discussed in later sections.

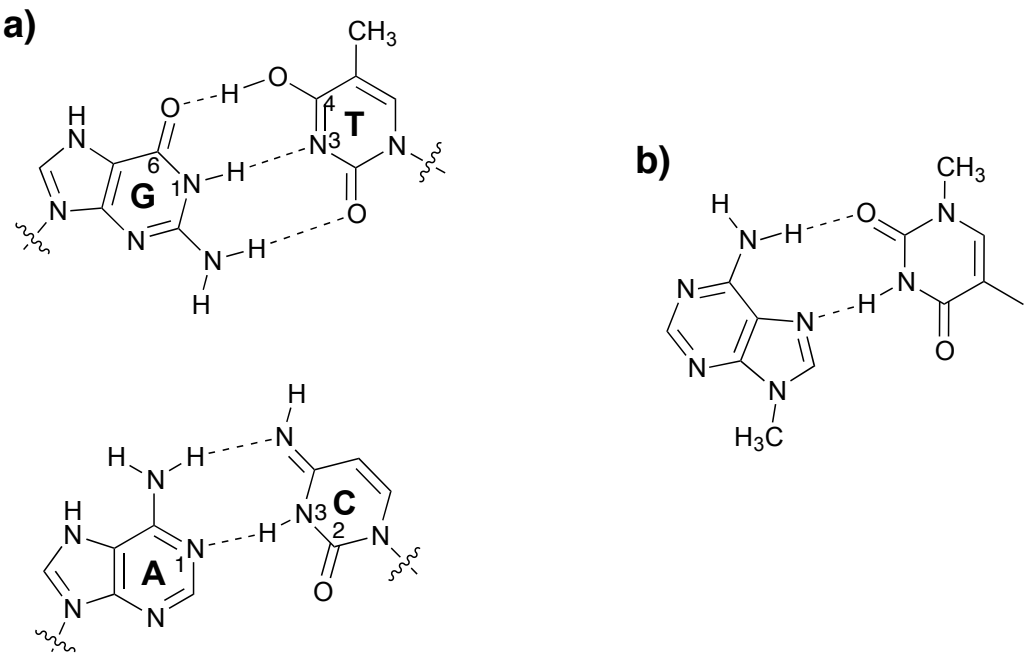


Figure 1.6 a) Tautomeric base pairing between guanine and thymine (top) and between adenine and cytosine (bottom). b) Hoogsteen base pairing between 9-methyladenine and 1-methylthymine.

1.2.4 X-ray crystal structure of DNA

Several problems impeded the unambiguous authentication of the structure of DNA at the time of Watson and Cricks' discovery. Direct observation techniques of biological molecules, such as X-ray crystallography were still in their infancy and required considerable time, effort and material to perform. No synthetic methodologies yet existed that could prepare measurable quantities of short, custom-sequence oligonucleotides for crystallization. Hence, the true structure of DNA proved elusive for almost three decades until Dickerson and co-workers reported the first X-ray crystal structure of a B-DNA duplex featuring the self-complementary dodecameric sequence d(CGCGAATTCGCG) that includes the six-base pair target sequence (GAATTC) of EcoRI, a restriction endonuclease (Figure 1.7).[24, 25] The structure was found to be

identical to the Watson-Crick model of B-DNA with the sole exception of base pair orientation, showed to be twisted slightly off-planar in a propeller-like configuration.

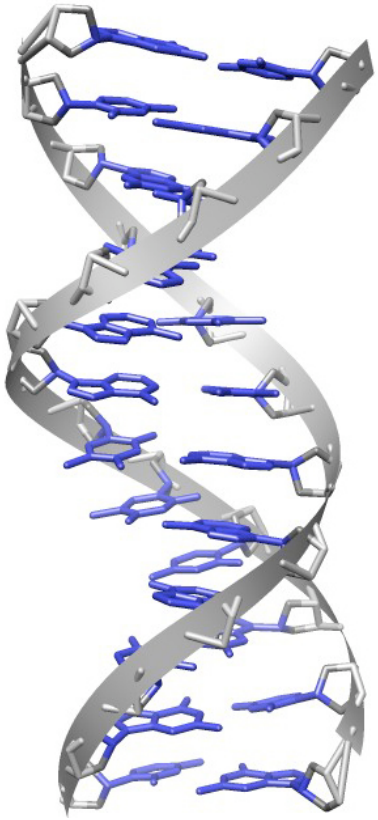


Figure 1.7 X-ray crystal structure of DNA dodecamer d(CGCGAATTCGCG) (PDB: 1BNA) showing the double helix. Visualization rendered with USCF Chimera molecular modeling software.[26]

1.3 Supramolecular nucleobase assemblies in DNA

1.3.1 G-quartets and quadruplexes

The propensity of guanine to self-assemble is well-known, owing to its π-electron rich molecular surface and terminally located dipoles encouraging stacking and planar end-on

interactions. Empirical evidence of aggregation *in vitro* was first detected in the form of aqueous gels where solutions of guanosine exhibited this unusual property in the presence of K^+.[27] In 1962, Gellert and co-workers reported analogous gel-forming activity of guanosine monophosphate (5'-GMP) at varying pH values and concentrations in the presence of sodium salt.[28], while K^+ and Sr^{2+} were later found to also induce 5'-GMP gelling with higher thermal stability.[29] It was proposed that the GMP residues self-assemble into planar tetrads supported by a network of hydrogen bonds, attaining higher order structure by stacking atop one another to form quadruplexes. In this model, tetrad formation involves a series of eight Hoogsteen hydrogen bonds. Each guanine base contributes a pair of hydrogen bond acceptors (O6 and N7) and a pair of donors (amino H and H1) that are positioned almost 90º relative to one another, giving rise to a four-fold centrosymmetric system (Figure 1.8a). The cation lies within the central cavity, coordinated by O6 of each base. Quadruplex formation ensues with individual quartets engaging in a helical π-stacked fashion approximately 3.25 Å apart with cation positioning dependent on its size (Figure 1.8b). Thermodynamic studies and *in silico* modeling have determined cation selectivities in the order $Sr^{2+}>K^+>Na^+\geq Rb^+>Li^+>Cs^+$.[30] Apart from size and charge density, the energies of ligand coordination and hydration are implicated.[31, 32]

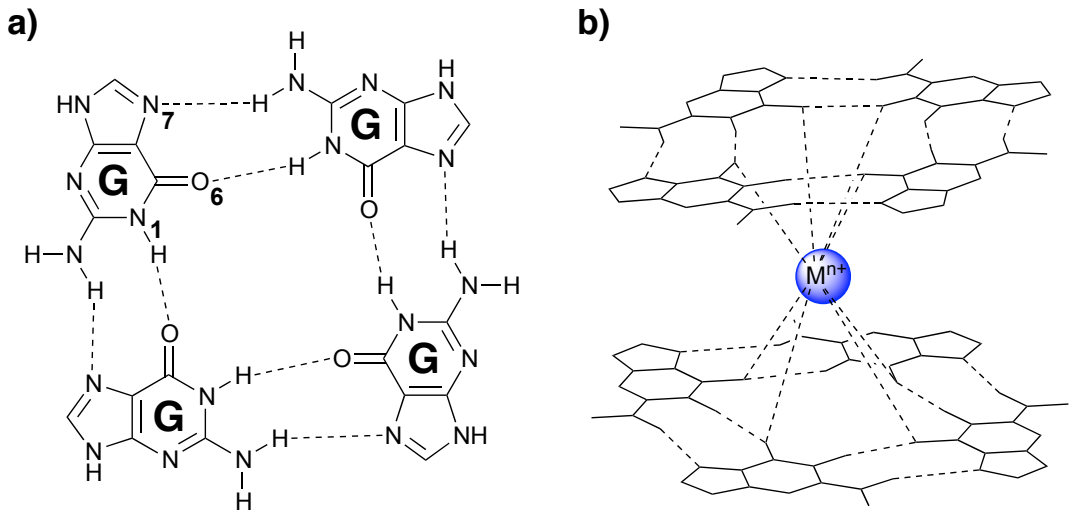


Figure 1.8 a) Hoogsteen hydrogen-bonded G-quartet. b) Helical stack of G-quartets to form a G-quadruplex with a sandwiched central cation coordinated by the O6 atom of each guanine base in a square antiprismatic geometry.

Structural proof of G-quadruplexes was provided by X-ray diffraction studies on poly(G)-containing RNA fibers. Zimmerman and co-workers reported the observation of parallel tetra-stranded G-quadruplexes (Figure 1.9a).[33] The data supported Gellert's planar, fourfold-symmetric quartet model and concerted hydrogen bonding mode, as well as the stacked quadruplex assembly. Geometric measurements revealed a right-handed helicity of approximately 31º per quartet layer. Spatial constraints within this arrangement lead to the adoption of the C3'-*endo* puckering of the ribose sugar unlike that of ideal B-DNA. Subsequent studies on monomeric GMP co-crystallized with NaCl revealed that individual nucleotides unbound to phosphate-sugar backbones are likewise able to form quartet and quadruplex assemblies,[34] demonstrating their profound thermodynamic stability.

G-quadruplexes formed from single and double-stranded oligonucleotides may display fully parallel or both parallel and anti-parallel G repeats, interspaced by non-G linkers providing

for hairpin loop formation. Sen and Gilbert characterized an intramolecular G-quadruplex stabilized at physiological salt concentrations with all four component poly(G) strands parallel,[35] similar to Zimmerman's system, while Williamson et al. reported an unusual tandem parallel/anti-parallel structure formed made possible by the appropriate *syn*/*anti* conformational switches of the guanine bases in each tetrad (Figure 1.9b).[36] Sundquist and Klug observed an anti-parallel bimolecular system (Figure 1.9c).[37] Subsequent investigations led by Sen probed the role of cations in determining the directional sense of G-quadruplex assemblies.[38]

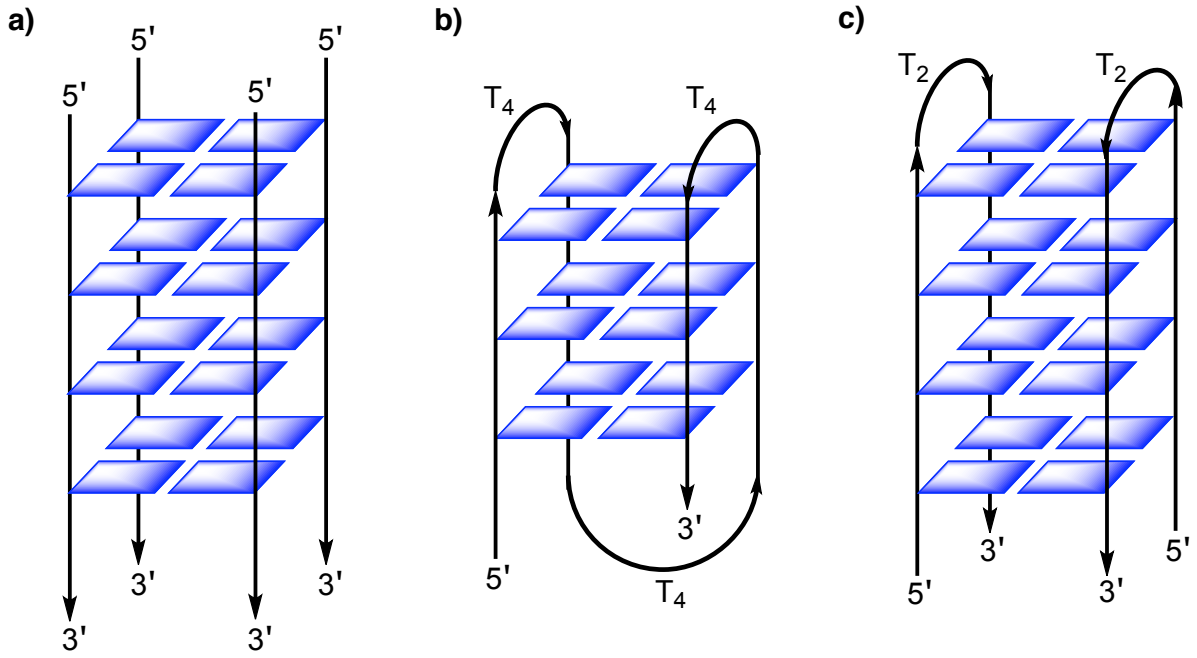


Figure 1.9 a) Cartoon representations of several G-quadruplex patterns. a) All-parallel system derived from poly(G) strands reported by Zimmerman. b) Intramolecular antiparallel system of d[$(T_4G_4)_4$] reported by Williamson. c) Sundquist's bimolecular antiparallel system adopted by d$(T_2G_4T_2G_4)_2$ strands. Guanine bases are represented by blue slabs.

1.3.1.1 Telomeric G-quadruplexes

In the "end replication problem",[39-41] repeated replication of somatic cells results in the gradual truncation of chromosomal DNA at their telomeric end regions due to the inability of DNA polymerase to synthesize the terminal 5' end of the lagging strand without an upstream 3'-hydroxyl group, leaving a characteristic overhanging flap at the 3' end of the leading strand. Over time, this leads to cell senescence and eventually apoptosis unless the excised portion is restored by the action of telomerase. Overexpression of this enzyme results in the unregulated repair of the telomeres, thus immortalizing the cell and ultimately leading to the uncontrolled cell replication and growth characteristic of cancer. The G-richness of the telomeric strand, especially the 3' overhanging region of between 100 – 200 bases in length,[42, 43] confers it a high propensity for G-quadruplex formation, thereby implicating these structures in telomerase activity regulation.[44] This has led to the design of ligands capable of binding and stabilizing G-quadruplexes as potential anticancer drugs.[45-47] The inhibition of DNA helicases involved in recombination by these ligand-quadruplex adducts further evidences the crucial role played by G-quadruplex stability in regulating cell lifespan and replication.[48]

Prior to the characterization of human telomeres, studies by Blackburn identified the 6-bp telomeric repeat sequence d(TTGGGG) of *Tetrahymena*,[49] although it was still uncertain if the leading or lagging strand was G-rich. The realization that telomerase is in fact a ribonucleoprotein functioning as a reverse transcriptase helped answer that question and address the end replication problem.[50, 51] The RNA component of *Tetrahymena* telomerase was found to contain the sequence r(AACCCCAAC) allowing targeting of the 3' overhang of the leading strand via a 3-bp, codon-like complementary base-pairing. Once in place, polymerase function

proceeds to fill in the remaining positions corresponding to a repeat unit. Subsequent cycles of translocation and polymerization gradually replenish the G-rich leading strand, now positively identified as the telomeric strand. Elongation of the C-rich lagging strand is then carried out by the action of DNA polymerase via Okazaki fragments.

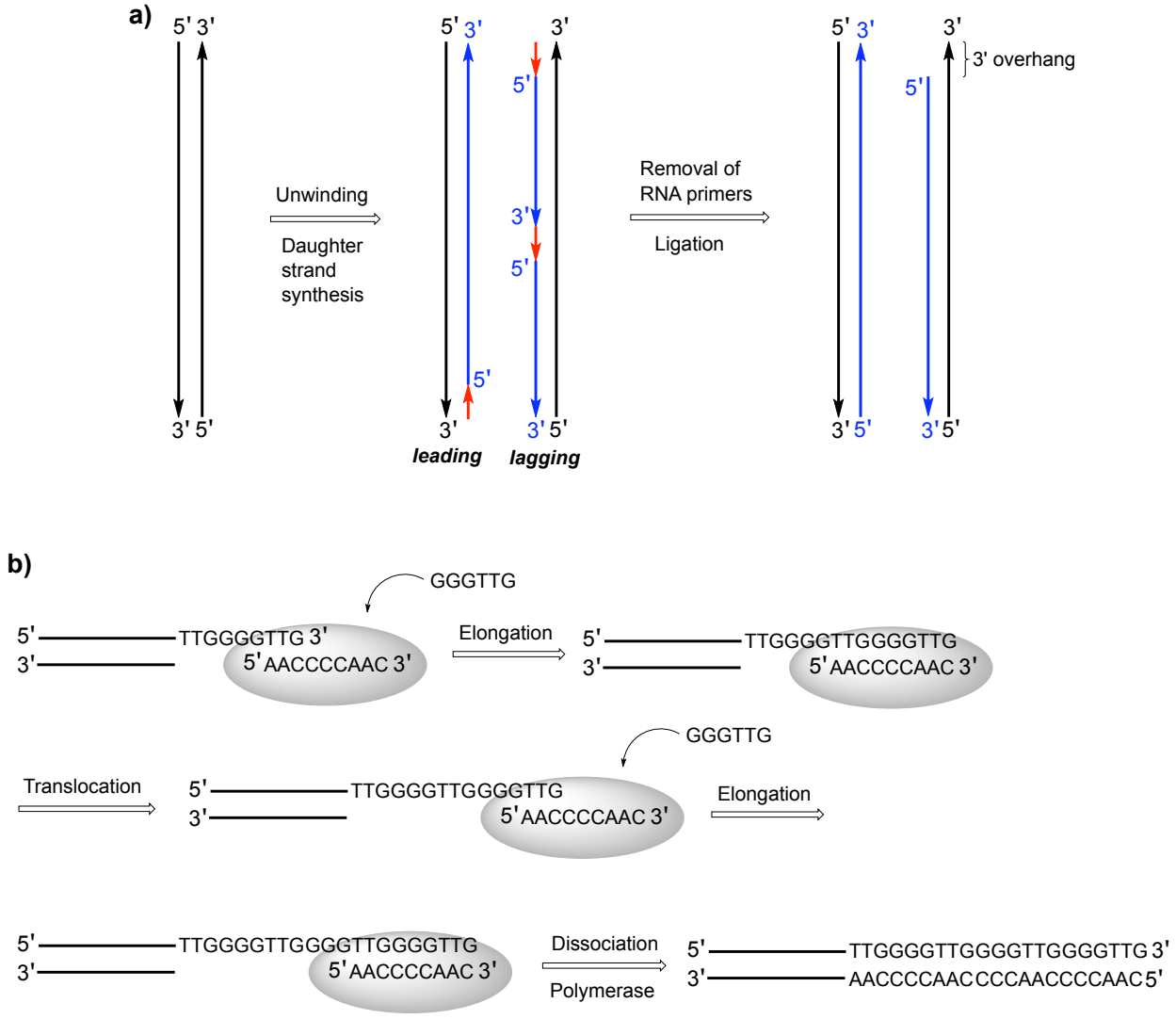


Figure 1.10 a) Schematic of DNA replication resulting in a 3' overhanging region of the lagging parent strand. b) Schematic of *Tetrahymena* telomere repair by the reverse transcriptase mechanism of telomerase.

In 1988, Moyzis and co-workers successfully elucidated and cloned the human telomeric repeat d(TTAGGG) which spans over a range of approximately five to eight kilobases.[52] Studies in which the fluorescent-labelled telomeric chimera d[$(GGGTTA)_7$•$(TAACCC)_7$] was exposed to the chromosomal material of 91 extant eukaryotic species including humans found their unanimous incorporation into the telomeres, suggesting that the sequences are highly conserved throughout all vertebrates thought to be descended from a common ancestor.[53] As with *Tetrahymena*, telomere repair in humans is mediated by a ribonucleoprotein-based telomerase, differing only in respect to the RNA recognition sequence being the eleven base-pair r(CUAACCCUAAC).[54]

Discrete telomeric repeat sequences have been shown to spontaneously adopt G-quadruplexes in both the solid and solution states. Kang and co-workers solved the first high resolution X-ray crystal structure of the *Oxytricha* sequence d(GGGGTTTTGGGG) in the presence of K^+ to reveal a bimolecular, antiparallel G-quadruplex with the two strands adjacent and the poly(T) spacers constituting hairpin loops extending laterally out of the quadruplex core (Figure 1.11a).[55, 56] Individual bases in each strand were found to alternate between the *syn* and *anti* glycosidic bond conformations. Interestingly, with Na^+ in the solution state, a remarkably different over-and-under arrangement is adopted with the two hairpin loops diagonal, as characterized via NMR (Figure 1.11b).[57] This conformational switching was attributed to the cation effect.[58] The crystal structure of the *Tetrahymena* telomeric repeat d(TGGGGT) was subsequently elucidated by Laughlan and co-workers. The resultant G-quadruplex from this sequence features a tetra-stranded fully parallel framework with all bases in the *anti*

conformation. Stabilizing Na^+ ions are positioned coplanar with each G-quartet layer (Figure 1.11c).[59]

Wang and Patel provided the first structural analysis of the 22-mer human telomeric repeat d[AGGG(TTAGGG)$_3$] in solution via NMR and *in silico* molecular dynamics simulations.[60] The intramolecular Na^+-bound G-quadruplex displays alternating parallel and antiparallel poly(G) runs with sequential lateral and diagonal hairpin loops. Nearly a decade later, Parkinson and co-workers presented the first crystal structures of the 22-mer repeat as well as the shorter 12-mer sequence d($TAG_3T_2AG_3T$) in the presence of K^+ (Figure 1.11d).[61] Contrary to expectations, both the resulting intra- and intermolecular G-quadruplexes were found to be fully parallel, maintained by the hairpin loops linking the top of one poly(G) strand to the bottom of another. Consequently, the loops protrude out the sides of the vertical quadruplex scaffold in a propeller-like configuration. All the guanine bases were found to be in the *anti* conformation with C2'-*endo* sugar puckering in effect throughout. Consistent with earlier studies, the K^+ ions are sandwiched between G-quartet layers and coordinated to the eight carbonyl O6 atoms in the square antiprismatic geometry.

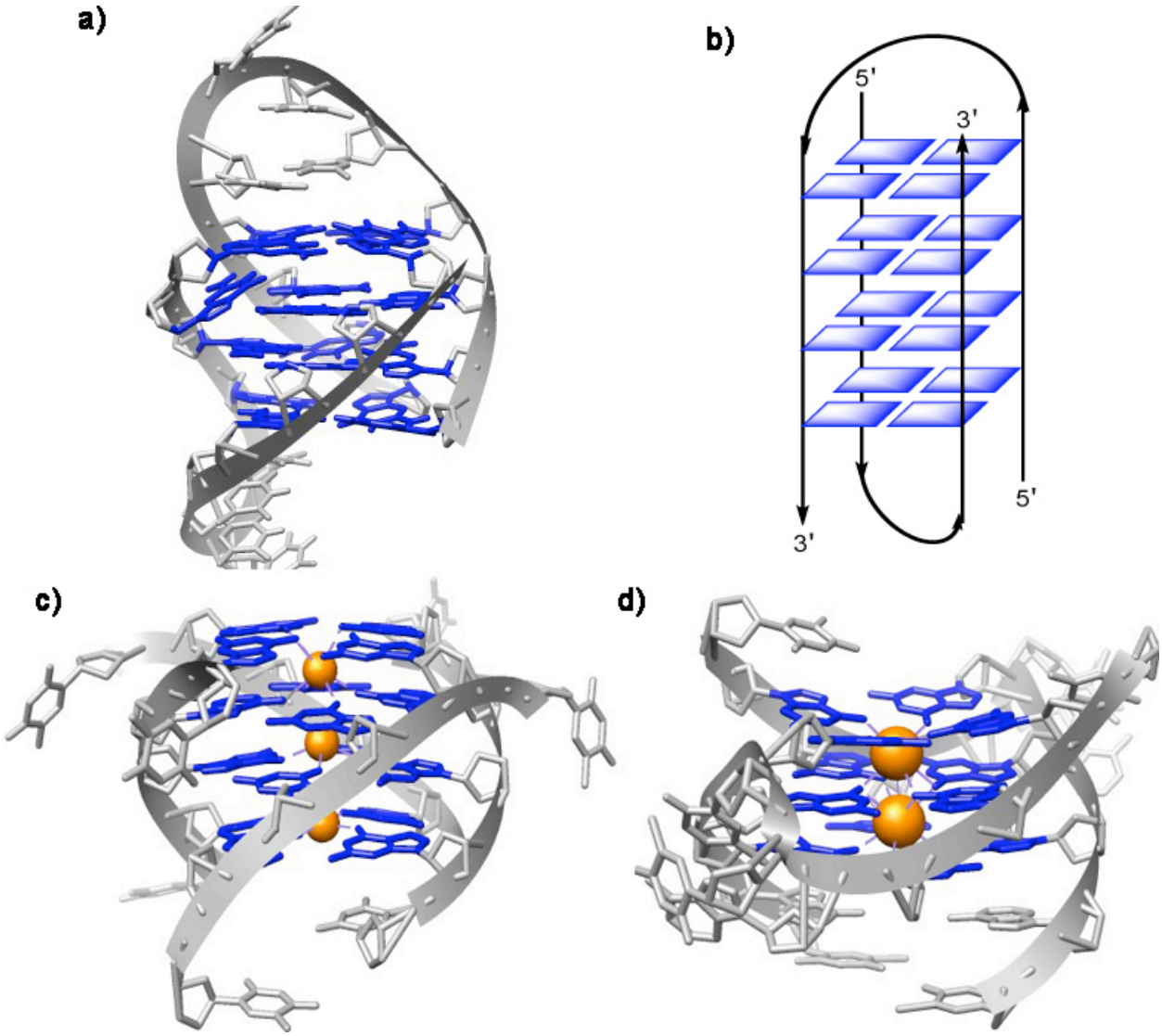


Figure 1.11 a) X-ray crystal structure of *Oxytricha* telomeric sequence d($G_4T_4G_4$) (PDB: 1D59). b) Schematic of diagonal bimolecular over-and-under G-quadruplex adopted by *Oxytricha* sequence in solution with Na^+. c) X-ray crystal structure of *Tetrahymena* telomeric repeat d(TG_4T) with sandwiched Na^+ (PDB: 244D). d) X-ray crystal structure of shortened human telomeric repeat d(T TAG_3TTAG_3T) with sandwiched K^+ (PDB: 1K8P). Crystal structures rendered with UCSF Chimera.

1.3.2 The i-motif

The structure-forming potential of C-rich sequences was recognized not long after the discovery of G-quadruplexes. In 1963, Langridge and Rich proposed from X-ray fiber diffraction experiments the possibility of poly(C) oligonucleotides adopting duplexes held together by hemiprotonated C•CH^+ base pairs at low pH (Figure 1.12a).[62] This novel base pairing was later found to be present in the X-ray crystal structure of acetylcytosine,[63] and subsequently in solutions of poly(C) RNA[64, 65] and DNA.[66] Nearly three decades later, NMR studies by Gehring ascertained the solution structure of the so-called i-motif adopted by the sequence d(TC_5) as a tetraplex featuring an intricately intercalated 'zipper' of C•CH^+ base pairs provided by two component duplexes (Figure 1.12b).[67] Both strands within each duplex were shown to be parallel, and both duplexes are oriented anti-parallel respect to each other. Right-handed helicity confers a twist of roughly 16º between base pairs on each duplex while cross-duplex base pairs are positioned orthogonally. The helical axis runs through the middle of all the base pairs.

The first crystal structure of an i-motif was delivered by Chen and co-workers of d(C_4),[68] revealing the expected intercalated C•CH^+ base pairs and directional sense of the strands (Figure 1.13a and b). Consistent with the solution structure, the helical twist was ascertained to be right-handed with an approximate 12.4º offset between layers. All bases were found to be in the *anti* conformation and sugar puckering was variable depending on their positions within the molecule, with the C4'-*exo* configuration being in majority. This allows some degree of structural freedom within the framework. Interestingly, base stacking was found to involve only the exocyclic heteroatoms (O2 and N4) as auxiliaries of the greater π-ring system, resulting in a

3.1 Å helical rise as opposed to 3.4 Å for B-DNA. Similar geometries have been documented in other pyrimidine-based structures.[69]

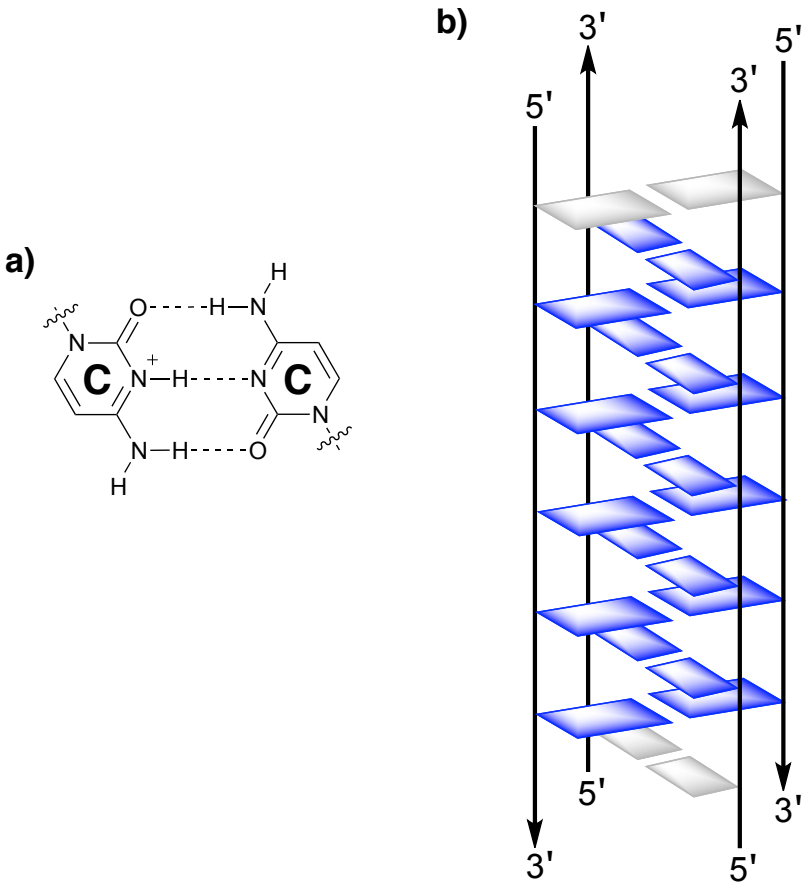


Figure 1.12 a) A hemiprotonated C•CH$^+$ base pair, the building block of i-motif architectures. b) Example of an i-motif: Schematic of an alternating parallel/antiparallel tetra-stranded structure adopted by d(TC$_5$) showing the intercalated hemiprotonated cytosine base-pairs. Blue and grey slabs represent cytosine and thymine bases respectively.

In 1995, the crystal structure of d(TAACCC), the vertebrate C-rich telomeric repeat was reported by Kang and co-workers,[70] describing a largely analogous intercalated C•CH$^+$ region to a previous study on d(CCCT) (Figure 1.13c).[71] However, helical rotation is inhomogeneous throughout, with an average of 19.7º. The C3'-*endo* sugar pucker features predominantly.

Interestingly, an intramolecular Hoogsteen A•T base pair exists in each strand, greatly stabilizing the overhanging hairpin loops. In addition, intermolecular Watson-Crick A•T base pairs between duplexes serve to lock the overall structure in place. Five years later, the solution structure of the extended 22-mer sequence d[$(CCCTAA)_3CCCT$] was solved by Phan and colleagues.[72, 73] The intramolecular i-motif identified presented a distinct major and minor groove bridged by the 3-bp AAT linkers.

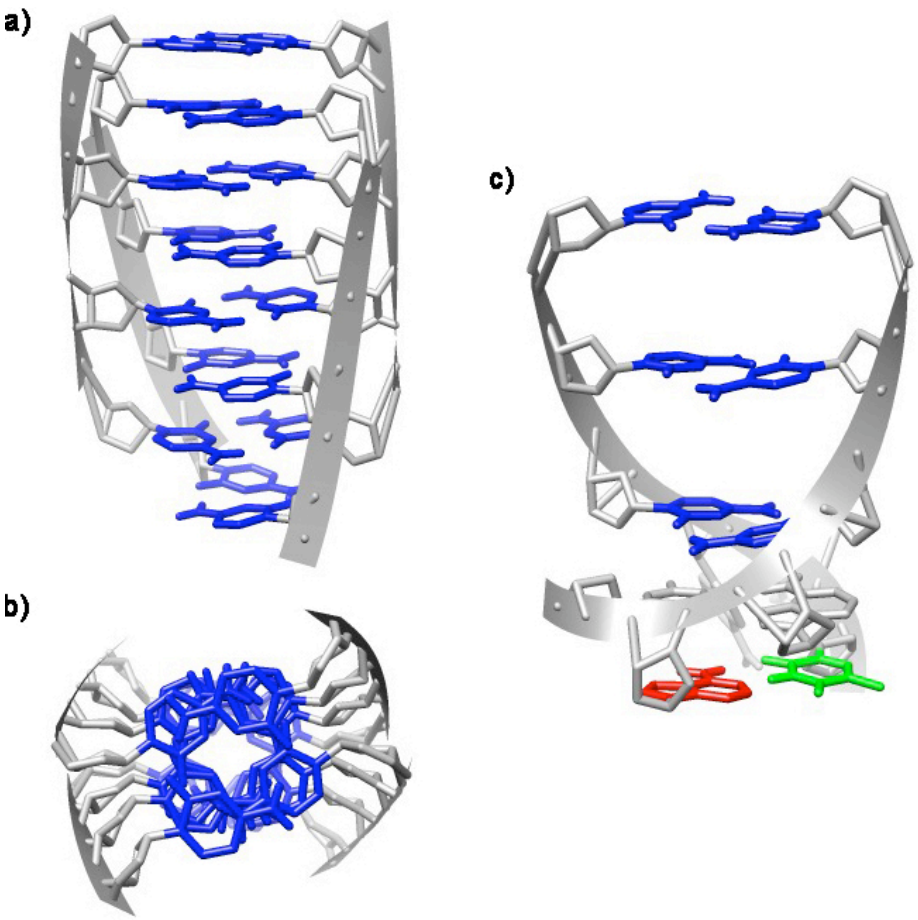


Figure 1.13 a) X-ray crystal structure of all-parallel d(C_4) i-motif tetraplex (PDB: 190D). b) Ditto. View down helical axis. c) X-ray crystal structure of bimolecular i-motif derived from d(TAACCC) (PDB: 200D). Intermolecular A•T base pair is represented by red adenine and green thymine residues. Crystal structures rendered with UCSF Chimera.

An anomalous i-motif assembled from the *Tetrahymena* telomeric repeat d(AACCCC) features unusual terminal adenine clusters made out of A•A base pairing between residues of adjacent strands.[74] The hydrogen bonding mode involves novel amino to N7 interactions giving rise to a rather complicated tetraplex in which the poly(C) regions extend out of an interwoven central adenine core (Figure 1.14).

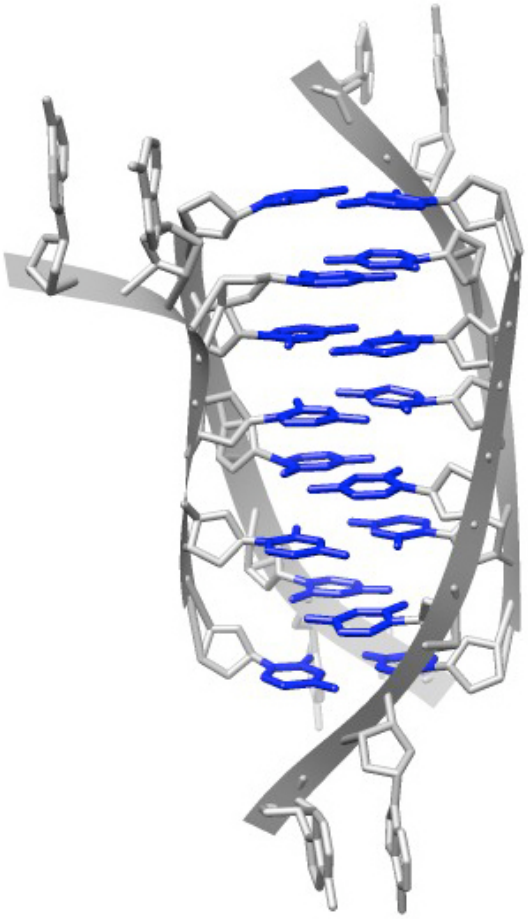


Figure 1.14 X-ray crystal structure of *Tetrahymena* telomeric repeat d(AACCCC) showing central i-motif core flanked by terminal adenine residues capable of intermolecular base-pairing (PDB: 294D). Rendered with UCSF Chimera.

1.3.3 Lipophilic G-quartets and quadruplexes

The biological significance of G-quadruplexes has prompted the preparation and study of synthetic analogous systems as model systems.[75] A number of these studies have entailed the derivatization of the guanosine subunits, largely affecting their solubility in water at physiological conditions and necessitating the use of organic solvents.

As discrete G-quartets are never found in nature due to their instability, their construction *in vitro* has been made possible only with the use of substrate control or template molecules. Sessler reported cation-free G-quartets assembled from an unnatural guanosine derivative in which the guanine moiety is blocked at the C8 position with a bulky dimethylaniline group (Figure 1.15a).[76] The added steric bulk drives the adoption of the *syn* configuration and in turn G-quartet formation. Under normal circumstances, guanosine monomers tend to aggregate into linear 'ribbons'. A guanosine-bound calix[4]arene dimer reported by Davis represents one of the pioneering examples of templated G-quartets (Figure 1.15b).[77] Each conjugate bore four guanosine residues covalently linked to a calix[4]arene macrocyle in a 1,3-alternate fashion. Quartet formation proceeds via end-on interaction between two guanosines of each subunit in the presence of either metal cations or in hydrated $CDCl_3$. This system was found to perform well as an ionophore for extracting cations into organic solvent. Several years later, Nikan and Sherman provided the first examples of fully lipophilic template-assembled synthetic G-quartets (TASQs) (Figure 1.15c).[78] In these systems, four guanosine subunits were covalently linked to the rims of rigid resorcinol-based cavitands bearing aliphatic pendant 'feet' to confer lipophilicity. Remarkably, in the absence of cations, the guanine bases spontaneously assemble into a quartet

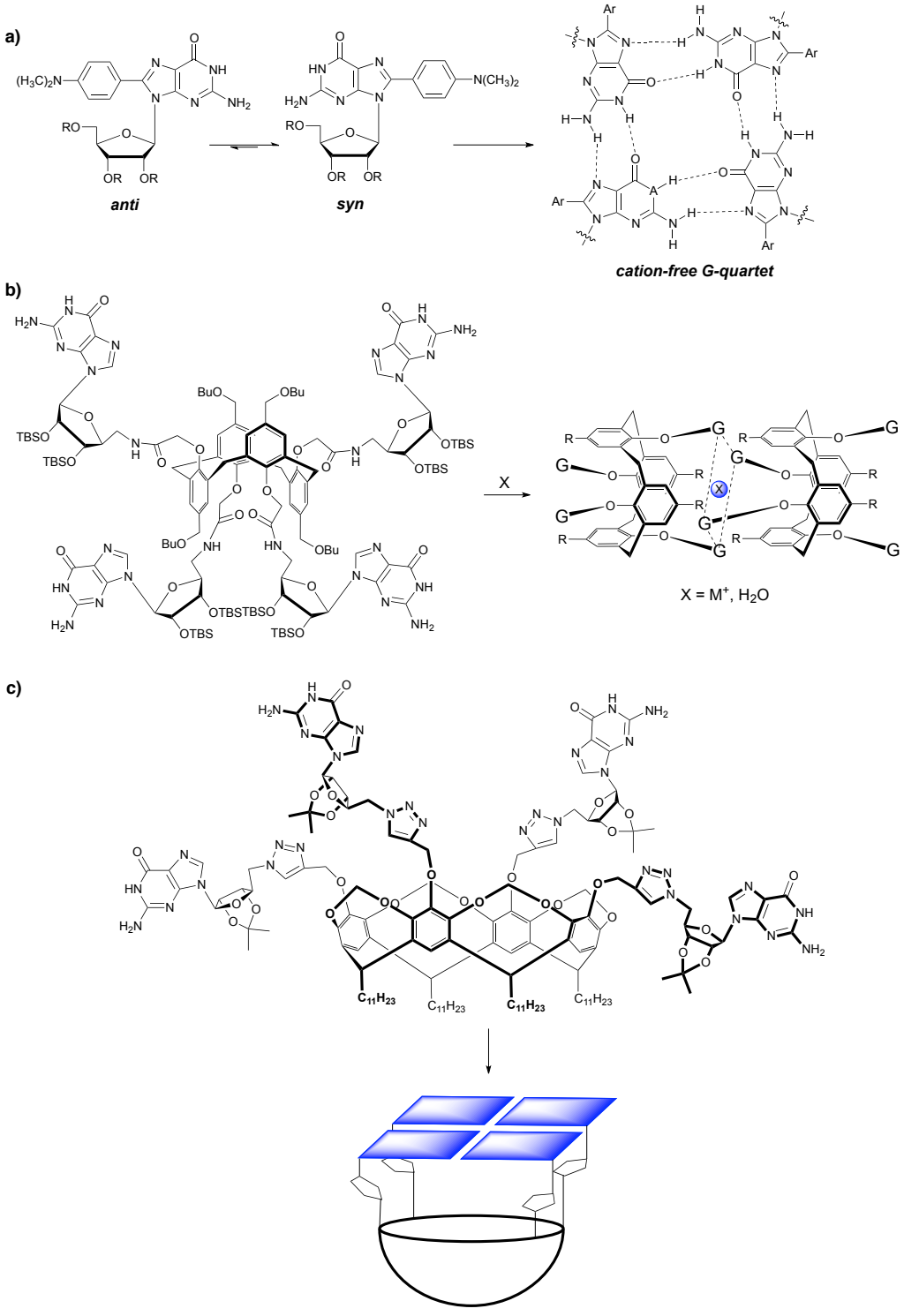


Figure 1.15 a) Cation-free G-quartet derived from *syn*-configured guanosine units. b) Bimolecular G-quartet supported by calix[4]arene templates. c) Cation-free G-quartet preorganized by a resorcinol-based cavitand.

in $CDCl_3$ at ambient conditions as evidenced by NMR. Subsequent cationic studies determined quartet assembly with Na^+, K^+ and Sr^{2+}, while interestingly Cs^+ affords a dimeric quadruplex.[79] More recently, Monchaud reported the assembly of G-quartets on EDTA-based[80] as well as porphyrin[81] templates. These 'DOTASQs' and 'PorphySQs' display strong G-quadruplex affinities that mimic natural ligand binding modes.

Synthetic lipophilic G-quadruplexes have been prepared from modified guanosine monomers. One of the earliest examples was reported by Davis and Gottarelli comprising two stacked G-quartets of 3', 5'-didecanoyl-2'-deoxyguanosine units, stabilized by a sandwiched, octacoordinated K^+ ion in the canonical square antiprismatic fashion.[82] A larger hexadecameric assembly involving four stacked quartets of 5'-silyl-2', 3'-*O*-isopropylidene guanosine units in the presence of K^+ and Cs^+ picrates was crystallized from $CDCl_3$.[83] The X-ray crystal structure shows three K^+ ions sandwiched between the quartet layers while a single Cs^+ ion rests atop the terminal quartet. Further stabilization is provided via novel hydrogen bonding between the guanosine amino protons uninvolved in Hoogsteen hydrogen bonding and the four picrate counteranions (Figure 1.16). Interestingly, the cations lie co-linearly along the central axis forming an ion channel. Its utility as an ionophore has been demonstrated. Further investigations showed that in the presence of divalent cations such as Sr^{2+} and Ba^{2+},[84, 85] spontaneous modification of the assembly occurs to provide a complex composed of two cation-stabilized octads. The reduced number of cations is attributed to their greater electrostatic repulsion. On the other hand, a more stable quadruplex is ensured with the increased positive charge density.

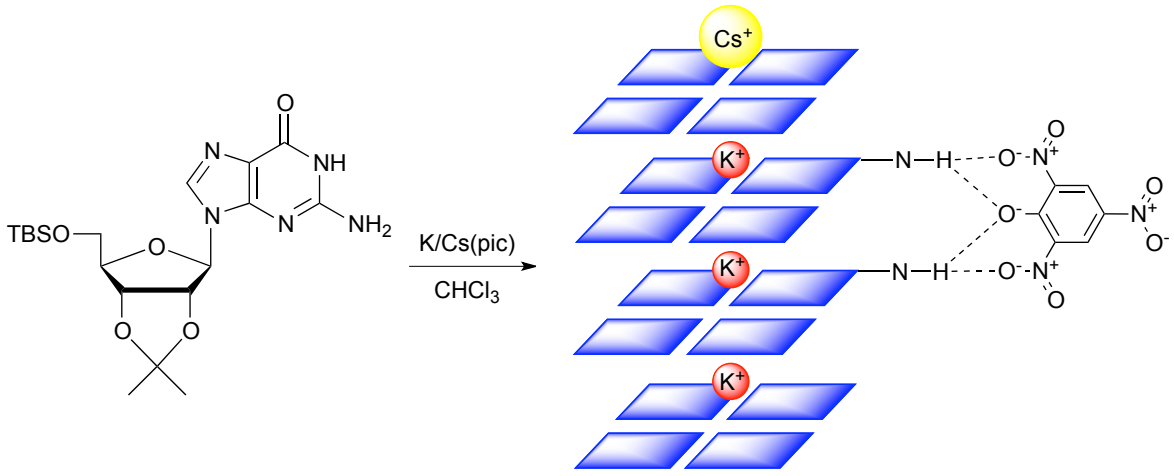


Figure 1.16 G-quadruplex assembled from lipophilic guanosine subunits in the presence of K and Cs picrates. Picrate 'clip' further stabilizes the system.

The crystal structure of a lipophilic G-quadruplex derived from the dimerization of two TASQs was reported by Nikan and Sherman.[86] Face-on dimerization of the G-quartets sandwiches a Na^+ ion in an encapsulated supramolecular system.

1.3.4 Pyrimidine-based quartets

The search for novel structures besides the ubiquitous guanine and cytosine-based motifs discussed so far is of importance because of their possible involvement in overall structure stability. Advances in analytical and computational techniques have allowed the identification and characterization of several unnatural quartet assemblies.

Uracil (U)-quartets were first reported by Cheong and Moore in an RNA quadruplex of sequence r(UGGGGU) in the solution state (Figure 1.17a).[87] NMR studies located the U-quartet at the 3' end of the all-parallel structure, assembled via imino H3•O4 hydrogen bond contacts

between each of the uracil bases (Figure 1.17b). This arrangement positions the imino proton and H5 of adjacent bases within range for NOE interaction, resulting in a characteristic cross-peak on the 2D NOESY spectrum. Several *ab initio* studies have correlated this bonding mode to an energy minimum.[88-90]

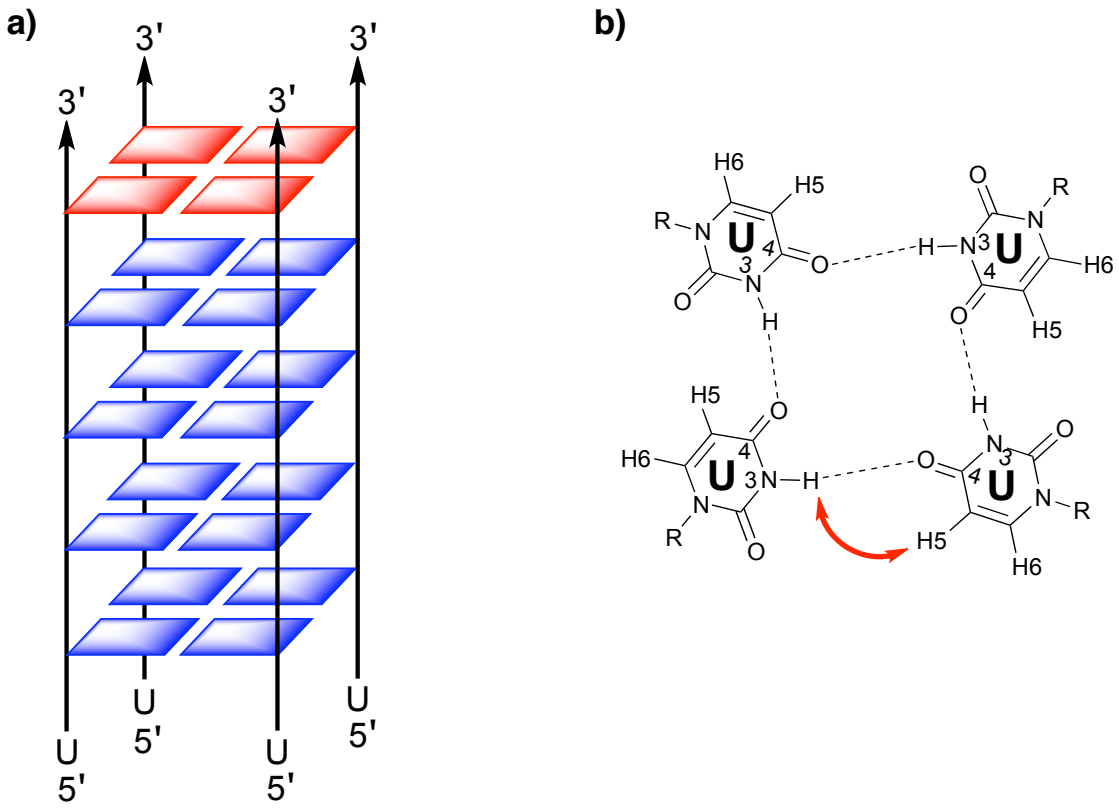


Figure 1.17 a) G-quadruplex scaffold (blue) capped by a 3' end U-quartet (red). b) U-quartet with imino to O4 hydrogen bonds. NOE interaction between imino and H5 of neighboring base is shown with a red arrow.

U-quartets have also been unambiguously identified in the solid state. Lippert reported the crystal structure of a 1-methyluracil/sodium tetrachloroaurate adduct showing the bases arranged in a planar quartet with the expected imino H3•O4 hydrogen bond framework.[91] Na^+ is positioned in-plane within the central cavity and is coordinated by the O4 atoms in a square planar geometry. A more recent report documents U-quartets formed out of 1-hexyluracil

subunits and stabilized in the presence of heavy, non-essential Ag^{+}.[92] The cations are located out of plane and coordinated to the O4 atoms in a square pyramidal fashion. Organization within the lattice is reminiscent of lipid bilayers with the hydrophobic n-hexyl chains buried exposing the U-quartets. Deng and Sundaralingam presented the first crystal structure of a U-quartet as part of a quadruplex derived from r(UGGGGU) in the presence of Sr^{2+}.[93] As with the solution structure,[87] the U-quartet was found at the 3' end. A notable difference, however, observes the uracil bases tapered inwards into the central cavity at an approximately 30º incline. Maintaining this conformation is the *anti* glycosidic conformation of the bases and C2'-*endo* sugar puckering. Sundaralingam subsequently observed a Na^{+}-bound U-quartet at the 3' end of a guanine-modified tetraplex r(U^{Br}dG)r(AGGU), where ^{Br}dG represents 8-bromodeoxyriboguanine.[94] Geometric properties of the quartet are similar with the Sr^{2+}-based system with further stabilization from water bridges connecting O2 on uracil and a phosphodiester oxygen atom.

In 2010, Xu et al. showed that the quadruplex $r(UAGGGU)_4$ containing a 3' end U-quartet is significantly more thermally stable than the modified sequence r(UUAGGG) without one.[95] Replacement of the uracil O4 atoms with sulfur resulted in structure destabilization, confirming not only the role of the U-quartet in overall quadruplex stability, but also the involvement of O4 in quartet formation.

Like uracil, thymine has been shown to exhibit quartet formation via the imino H3•O4 hydrogen bonding mode. Computational studies by Gu and Leszczynski revealed a higher stabilization energy for a non-planar T-quartet and assigned stronger hydrogen bonding for an off-planar, propeller twisted-type configuration compared to an ideal plane.[89] Adoption of this conformation is a consequence of the additional steric bulk introduced by the 5-Me groups.

Nevertheless, T-quartets have been observed, albeit incorporated in pre-existing G-quadruplex scaffolds. Patel and Hosur solved the NMR solution structure of quadruplex $d(TGGTGGC)_4$.[96] Observed NOE correlations between the imino H3s and 5-Me protons of the T4 residues confirmed the presence of the internal T-quartet. Molecular dynamics simulations predicted its conformation to be in agreement with Gu and Leszczynski's model. T1 residues were found to adopt a quartet at low pH (5.0) and temperature (5 ºC).

The crystal structure of $d(TGGGGT)_4$ co-crystallized with Na^+ and Tl^+ reported by Subirana et al. reveals a lattice consisting of stacked quadruplexes, their interfaces at which are located T-quartets stabilized by slightly off-planar Na^+ cations while Tl^+ is distributed throughout the G-quadruplex regions.[97] A single unit is shown in Figure 1.18. Consistent with previous findings the quartets assume the propeller-twisted configuration. More recently, Sket and Plavec solved the solution structure of the same quadruplex, revealing a T-quartet at the 5' end.[98] Interestingly, on replacing the 5' end thymines with uracils, a stacked quadruplex dimer ensues sandwiching a U-quartet, similar to that of the crystal structure reported by Deng and Sundaralingam for $r(UGGGGU)_4$.

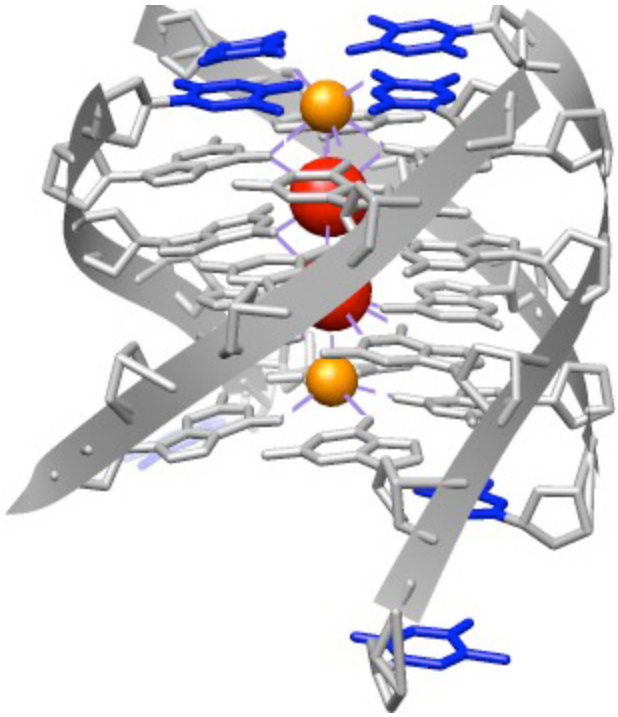


Figure 1.18 Crystal structure of d(TG$_4$T)$_4$ (PDB: 1S47) with top thymine quartet (blue) coordinating a Na^+ ion (orange) capping a G-quadruplex scaffold stabilized by Tl^{3+} ions (red). Rendered with UCSF Chimera.

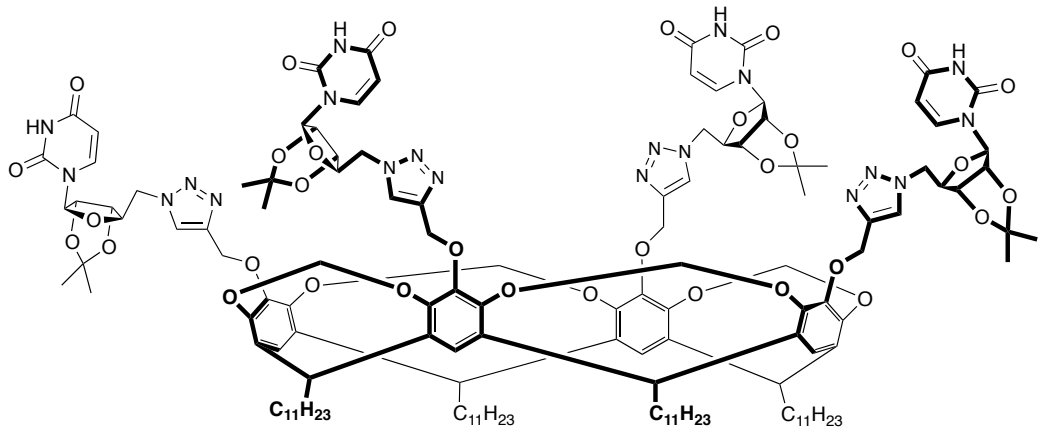
1.4 Thesis aim and goals

With ever-increasing evidence of the intrinsic role of hitherto poorly documented and understood uncommon nucleobase motifs, the impetus for this thesis arose from the standpoint of their *in vitro* construction via a synthetic organic approach for characterization and study. Our group has in recent years demonstrated the profound templating ability of resorcinol-based cavitands. With this in mind, could the substrate scope of these templates be expanded to include the rare pyrimidine tetrads like uracil and thymine? Could a degree of complexity be added by the installation of the corresponding quadruplexes? It is hoped that the synthesis and study of these systems might bestow on them more attention and shed more light on their topological significance.

Chapter 2: Synthesis and Characterization of a Template-Assembled Synthetic U-Quartet

2.1 Synopsis

The present study documents the design and synthesis of tetra-coupled uridine-cavitand conjugate **1** (Figure 2.1). Characterization was carried out using NMR and CD spectroscopic techniques to obtain evidence of the adoption of an unprecedented cation-free U-quartet by the pendant uracil moieties under lipophilic conditions. The system showcases the chemical versatility of cavitands to adapt to different solvent environments and accept a wide range of substrates. Their profound ability to preorganize discrete units of otherwise unstable nucleobase motifs with well-defined topology is also reflected. These assemblies may serve as models for thermal and kinetic stability studies.



1

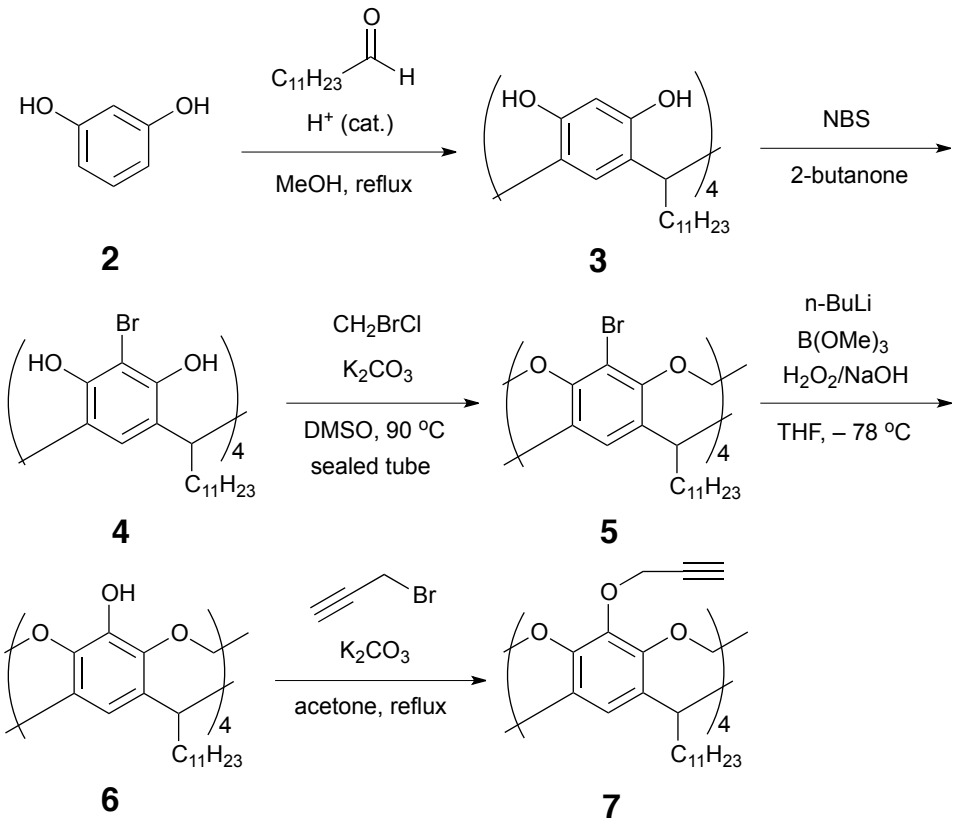
Figure 2.1 Side-on representation of conjugate 1.

2.2 CPK modeling studies

To assess the feasibility of a quartet assembly adopted by the four uracil bases of **1**, a space-filling Corey-Pauling-Koltun (CPK) model was constructed and examined to obtain information on conformational characteristics and pinpoint the presence of possible steric and electrostatic interactions. With glycosidic bond angles set to *anti*, the predominant conformation in native nucleic acids, the model revealed that a planar U-quartet could in fact be assembled over the rim of the cavitand architecture without torsional or steric strain. This arrangement also projects the hydrophobic isopropylidene groups installed on the ribose ring outwards into the extraneous environment, which may further promote its assembly under lipophilic conditions. The potential for ring stacking between the uracil and underlying triazole rings was also demonstrated, allowing for π-π interactions that may further stabilize the scaffold.

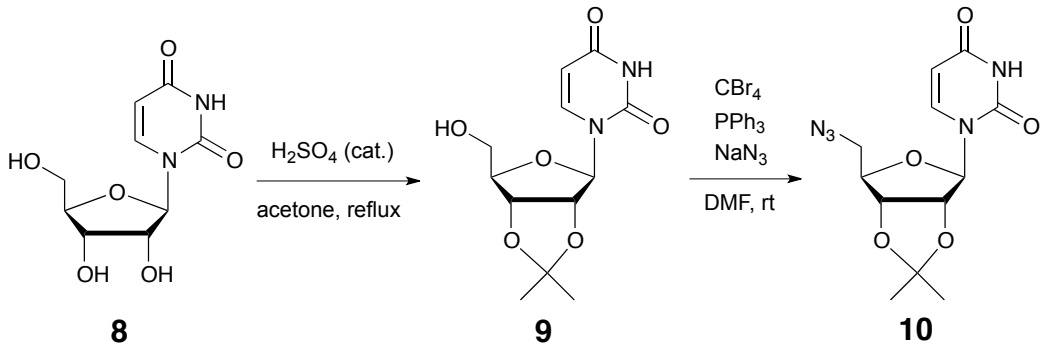
2.3 Synthetic strategy

Towards the synthesis of **1**, a convergent strategy was employed that involved first the preparation of cavitand **7** and 5'-azido-2', 3'-*O*-isopropylidene uridine **10**. Preparation of **7** was carried out in five steps according to literature procedures (Scheme 2.1).[99] Acid-catalyzed electrophilic aromatic substitution of resorcinol **2** with dodecanal provided resorcinarene octol **3** in quantitative yield. Treatment with *N*-bromosuccinimide (NBS) yielded the brominated derivative **4**, whereupon it was heated with bromochloromethane in the presence of K_2CO_3 as a base in a sealed pressure tube to afford the fully bridged cavitand **5** expediently.[100] Halogen-lithium exchange with *n*-butyllithium followed by the addition of trimethyl borate $B(OMe)_3$ and subsequent H_2O_2 oxidation gave tetrol **6**. Final treatment with propargyl bromide with K_2CO_3 base provided the propargylated cavitand template **7**.



Scheme 2.1 Synthesis of cavitand 7.

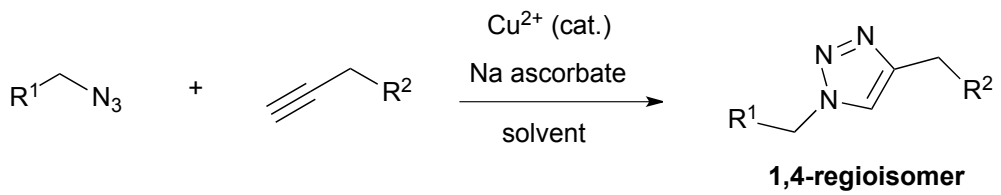
Synthesis of **10**[101] began with the acetalization of uridine **8** with acetone in the presence of mineral acid catalyst to afford 2',3'-*O*-isopropylidene derivative **9**. Treatment with CBr_4, PPh_3 and NaN_3 in a modified Mitsunobu procedure gave access to **10** in good yield (Scheme 2.2).



Scheme 2.2 Synthesis of 5'-azidouridine 10.

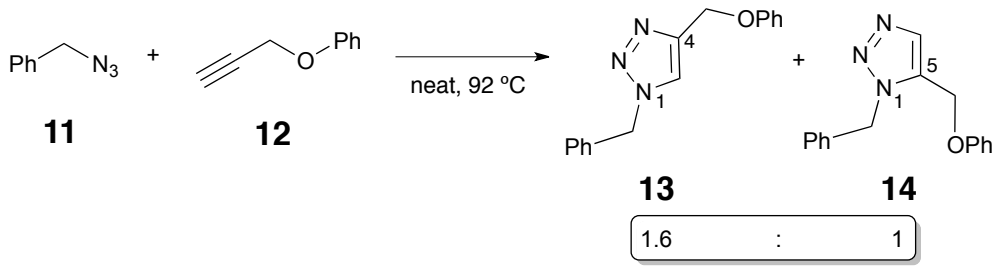
2.3.1 Conjugation via the "click" reaction

The Cu(I)-catalyzed azide-alkyne cycloaddition between azides and alkynes is the quintessential "click" reaction (Scheme 2.3) and provides a 1,4-disubstituted 1,2,3-triazole product exclusively. Developed by K. B. Sharpless in 2002,[102] the reaction involves the *in situ* reduction of an inorganic Cu(II) precatalyst by an appropriate reductant, most commonly sodium ascorbate. The resultant reactive Cu(I) species metalates the terminal alkyne position to form a Cu(I)-acetylide complex, which interacts with the azide reactant to initiate the catalytic cycle. The overall result is the coalescing of the azide and alkyne moieties to close the triazole ring. Density functional theory (DFT) studies have delivered a plausible mechanism for the reaction.[103]



Scheme 2.3 A general 1,4-regioselective "click" reaction.

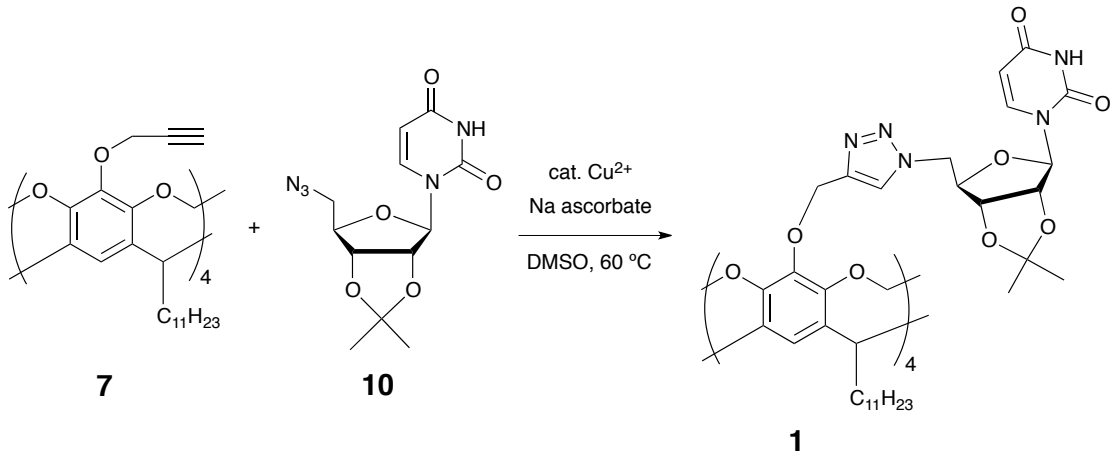
The forerunner of the "click" reaction is the Huisgen reaction, a thermally driven 1,3-dipolar cycloaddition between alkyl azides and terminal alkynes.[104] Pioneered by Rolf Huisgen, an early example is shown in Scheme 2.4 with the reaction between benzyl azide **11** and phenyl propargyl ether **12** to provide a mixture of 1,4 and 1,5-disubstituted triazoles **13** and **14** in close to a 1:1 ratio. The major shortcoming of this reaction is its lack of regioselectivity, which is circumvented by the Cu-catalyzed variant. The reaction therefore sees extensive use in bioconjugation where regioselectivity is imperative. Practical advantages include its robustness and versatility. It proceeds in a wide range of solvents and tolerates numerous functional groups. In addition, the reaction proceeds cleanly and is completely atom economical, facilitating work-up procedures.



Scheme 2.4 Example of a thermal Huisgen cycloaddition.

The key step for the synthesis of conjugate **1** employed the Cu(I)-catalyzed "click" reaction between the alkynyl and azide moieties of **7** and **10** respectively. The use of this reaction is desirable for several reasons: The length of the triazole backbone introduced between the nucleosides and the cavitand platform confers some measure of conformational flexibility. The π systems of the triazole and uracil rings could potentially participate in stabilizing stacking interactions.

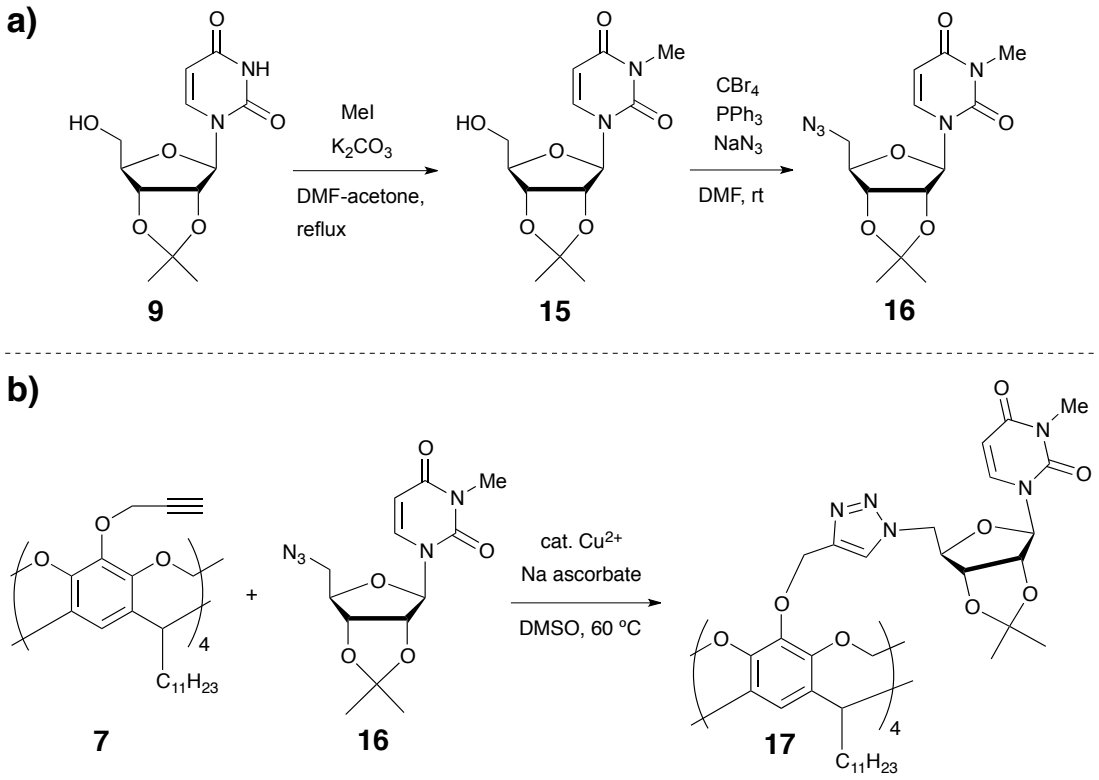
In the presence of a catalytic amount of copper (II) sulfate with sodium ascorbate as the reductant, the reaction between **7** and **10** proceeded smoothly under mild heating in DMSO to afford **1** in moderate yield (Scheme 2.5). MALDI-TOF mass spectrometry confirmed its identity as the tetra-coupled species. The compound was found to exhibit good solubility in chloroform and methanol.



Scheme 2.5 Synthesis of conjugate 1.

2.3.2 Synthesis of *N*-methylated conjugate

For subsequent control experiments, conjugate **17** was designed and prepared (Scheme 2.6). Methylation of all uracil imino positions in this molecule precludes hydrogen bonding, both intra- and intermolecular, thus preventing quartet assembly as well as higher-order aggregation. The synthetic procedure was identical to that of **1** with the exception of the extra *N*-methylation step of **9** to provide **15**, and finally **16**.



Scheme 2.6 Synthesis of a) azide 16 and b) conjugate 17.

2.4 NMR characterization studies

2.4.1 ^{1}H NMR signal assignment

Being symmetric and synthesized from enantiomerically pure starting materials, conjugates **1** and **17** display a single set of ^{1}H NMR resonances in both $CDCl_3$ (Figures 2.20 and 24) and DMSO (Figures 2.22 and 2.26). Signal assignment was performed with the use of ^{1}H and ^{1}H-^{1}H COSY NMR (spectra provided in the Experimental Section), with ^{1}H-^{1}H NOESY employed to reveal important spatial characteristics and regiochemistry. Conjugate **17** was assigned and characterized in all manners identical to **1**, and therefore will not be discussed in

the following sections. Figure 2.2 shows the condensed structure of **1** with all protons labeled. Important COSY correlations are illustrated in Figure 2.3.

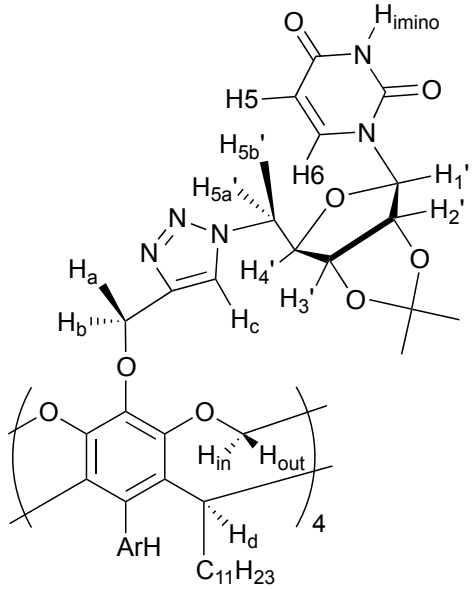


Figure 2.2 Condensed structure of 1 with protons labeled.

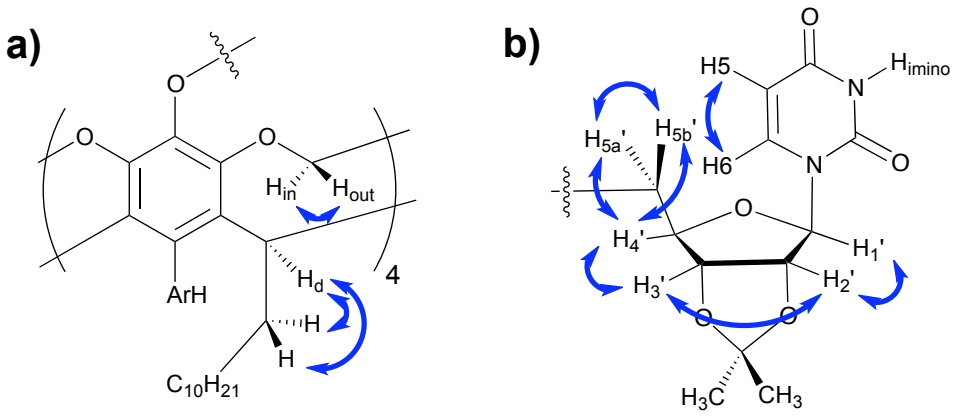


Figure 2.3 Expected COSY correlations in a) the cavitand moiety, between H_{in}/H_{out} and H_d/methylene 'foot', b) the nucleoside moiety, amongst ribose protons and H5/H6 on uracil.

The acidic H_{imino} proton is expected to have the largest chemical shift. Deuterium exchange with CD_3OD results in complete signal loss, confirming its identity (Figure 2.4). H5 and H6 are readily identified by their relatively large coupling constants ($^3J \sim 8.0$ Hz) typical of *cis*-alkene protons. H6 is located downfield of H5 due to its proximity to the electronegative N1 atom of uracil as well as the resonance effect of the enone.

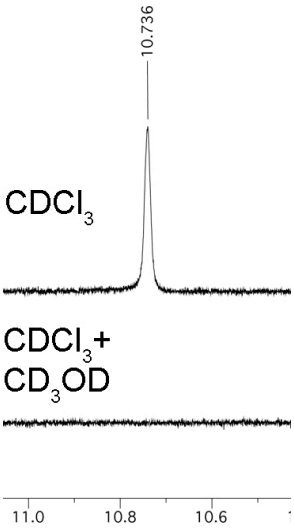


Figure 2.4 H_{imino} signal is quenched on exchange with CD_3OD.

The ribose ring protons follow the Karplus rule governing coupling constant magnitude and COSY crosspeak intensities.[105, 106] This is especially evident in the H_1'/H_2' and H_3'/H_4' interactions, where their approximately orthogonal dihedral angles cause attenuation or loss of signal intensity. In other cases non-orthogonal (H_4' to H_{5a}'/H_{5b}') and near-eclipsed (H_2' to H_3') protons give rise to clearly visible and intense crosspeaks respectively. Representative examples are shown in Figure 2.5. Assignment of this series of protons begins with H1', being most

deshielded and located furthest downfield. From this position, consecutively upfield assignments are made based on the appropriate COSY data.

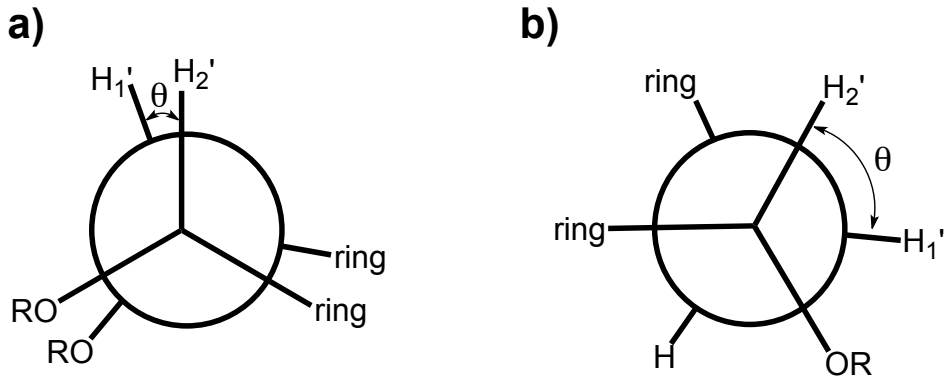


Figure 2.5 Dihedral angles of a) H_2' and H_3', eclipsed and therefore yielding a ^3J maximum and b) H_1' and H_2', close to orthogonal thus yielding a ^3J minimum. –OR groups represent the isopropylidene moiety.

Within the cavitand scaffold, geminal coupling between the upper bridge methylene protons (H_{in} and H_{out}) results in a fairly large coupling constant ($^2J \sim 7.0$ Hz) and strong COSY correlation. A large chemical shift difference separates them; H_{out} is located almost 2.0 ppm downfield of H_{in}, shielded by the anisotropic current of the cavitand as it extends inwards towards the cavitand cavity, while H_{out} projects outwards and is deshielded. The benzylic methine proton (H_d) on the lower cavitand bridge couples with its methylene neighbour on the long chain aliphatic 'foot'. The aryl proton (ArH) is assigned based on the typical chemical shift value of a benzene ring, although its proximity to H_d in space makes its unambiguous identification by ^{1}H-^{1}H NOESY possible (vide infra).

2.4.1.1 Signal assignment and regiochemical analysis by NOESY

^{1}H-^{1}H NOESY is a powerful tool that allows identification of proton correlations through space and is used extensively in structure elucidation and confirmation. As expected, the NOE spectrum of **1** reveals the correlation between ArH and H_d (Figure 2.6).

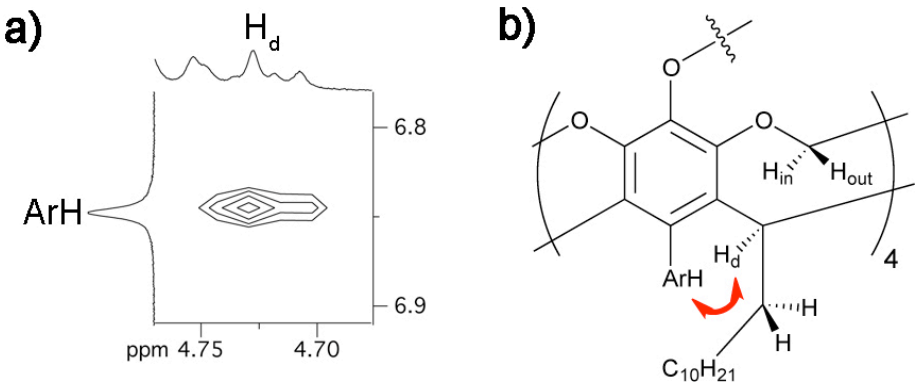


Figure 2.6 a) Portion of 400 MHz NOE spectrum of 1 in $CDCl_3$ showing crosspeak between ArH and H_d. b) NOE interaction between ArH and H_d.

The triazole proton (H_c) is easily identified by NOE correlations with both pairs of flanking methylene linkers (H_a/H_b, H_{5a}'/H_{5b}') (Figure 2.7). This also confirms the 1,4-regiochemistry of the triazole moiety expected of the regioselective Cu(I)-catalyzed "click" reaction, as the distance between H_c and H_{5a}'/H_{5b}' in the 1,5-regioisomer exceeds that required for minimal NOE interaction (~ 5 Å).

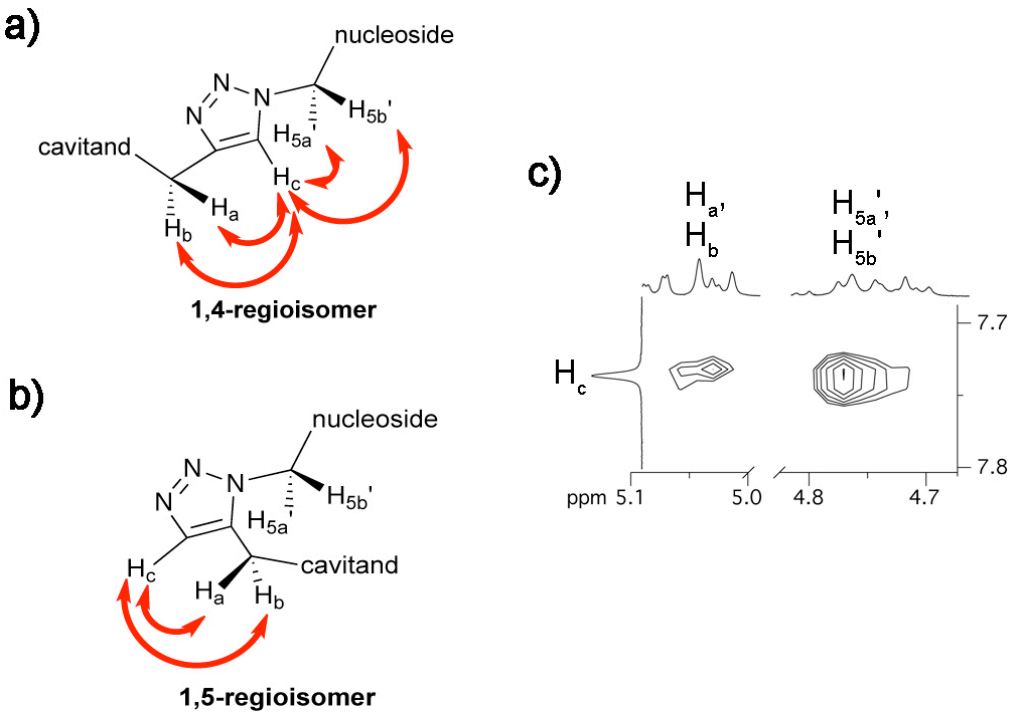


Figure 2.7 NOE interactions in the a) 1,4 and b) 1,5-regioisomers of 1. c) Portion of 400 MHz NOE spectrum of 1 in $CDCl_3$ showing NOE crosspeaks from H_c to H_{5a}'/H_{5b}' and H_a/H_b.

2.4.2 Summary of ^{1}H signal assignments

Remaining protons to be assigned include those of the isopropylidene methyl groups on the ribose sugar as well as the long aliphatic chain cavitand 'feet'. These protons are located in the typical alkyl chemical shift region. Tables 2.1 and 2.2 lists all protons in **1** and **17** with the corresponding COSY and/or NOESY data used in their assignment.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
H_{imino}	10.74	11.46		
H5	5.71	5.61	H6	
H6	6.98	7.64	H5	
H_1'	5.62	5.77	H_2'*	
H_2'	5.08	5.12	H_1'*, H_3'	
H_3'	5.04	4.89	H_2', H_4'**	
H_4'	4.58	4.40	H_3'**, H_{5a}'/H_{5b}'	
H_{5a}'	4.75	4.75	H_4', H_{5b}'	
H_{5b}'	4.75	4.65	H_4', H_{5a}'	
H_a	5.05	4.94		H_c
H_b	5.16	4.94		H_c
H_c	7.74	8.20		H_a/H_b, H_{5a}'/H_{5b}'
H_d	4.77	4.58	-CH_2- 'feet'	ArH
H_{in}	4.37	4.28	H_{out}	
H_{out}	5.79	5.93	H_{in}	
ArH	6.85	7.29		H_d
-CH_2- 'feet'	2.20	2.33	H_d	
long chain 'feet'	1.29 – 1.38	1.24 – 1.30		
terminal CH_3	0.92	0.85		
isoprop. CH_3	1.29	1.29		

* Expected but unobservable COSY signals.

** Very weak COSY signals.

Table 2.1 ^{1}H signal assignments of 1 in $CDCl_3$ and DMSO-d_6 at 25 ºC.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.30	3.15		
H5	5.74	5.75	H6	
H6	7.18	7.70	H5	
H_1'	5.60	5.82	H_2'*	
H_2'	5.04	5.13	H_1'*, H_3'	
H_3'	5.04	4.94	H_2', H_4'*	
H_4'	4.58	4.42	H_3'*, H_{5a}'/H_{5b}'**	
H_{5a}'	4.67	4.76	H_4'**, H_{5b}'	
H_{5b}'	4.77	4.64	H_4'**, H_{5a}'	
H_a	5.05	4.94		H_c
H_b	5.05	4.94		H_c
H_c	7.67	8.20		H_a/H_b, H_{5a}'/H_{5b}'
H_d	4.67	4.58	-CH_2- 'feet'	ArH
H_{in}	4.32	4.28	H_{out}	
H_{out}	5.74	5.91	H_{in}	
ArH	6.80	7.30		H_d
-CH_2- 'feet'	2.17	2.33	H_d	
long chain 'feet', isoprop. CH_3	1.25 – 1.28	1.25 – 1.28		
terminal CH_3	0.88	0.86		
isoprop. CH_3	1.54	1.47		

* Expected but unobservable COSY signals.

** Very weak COSY signals.

Table 2.2 ^{1}H signal assignments of 17 in $CDCl_3$ and DMSO-d_6 at 25 ºC.

2.4.3 NMR solution structure

The ^{1}H NMR spectra of **1** show a significant downfield shift (~ 1 ppm) of H_{imino} in DMSO-d_6 (Figure 2.22) relative to $CDCl_3$ (Figure 2.20), indicating a greater deshielded state therein. This is expected, as DMSO is a strong hydrogen bond acceptor. Any hydrogen-bonded assembly adopted by **1** in $CDCl_3$ is therefore destabilized in DMSO. In order to obtain more evidence of structure in **1** in $CDCl_3$, the ^{1}H NMR spectrum of **1** was compared with that of free nucleoside **10** (Figure 2.8).

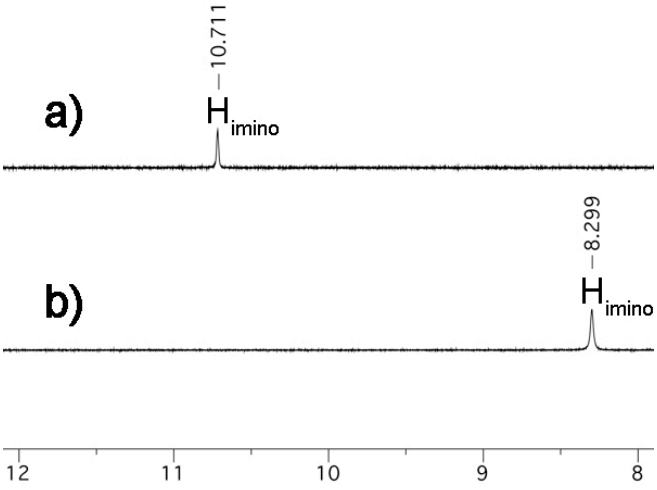


Figure 2.8 Portions of 400 MHz ^{1}H NMR spectra of a) 1 and b) 10 in $CDCl_3$ at 25 ºC.

A significant downfield shift from 8.3 to 10.7 ppm ($\Delta_{\delta H}$ ~ 2.5 ppm) for H_{imino} occurs on attachment of the uracil moieties to the cavitand, indicating deshielding of these protons likely caused by hydrogen bonding in structure assembly. To probe the nature of this structure, NOESY was performed which revealed a crosspeak between H_{imino} and H5 (Figure 2.9a), the key correlation used to authenticate U-quartets.[87] The low intensity of this crosspeak suggests a

loosely coordinated U-quartet. On lowering the temperature to –20 ºC, H_{imino} shifts further downfield to 11.5 ppm, close to that reported for the imino proton of guanine residues in the TASQs characterized by the Sherman group,[78] thereby indicating a greater degree of hydrogen bonding and a more tightly associated quartet. The increased intensity of the H_{imino}/H5 crosspeak further supports this notion (Figure 2.9b). This crosspeak is absent in free nucleoside **10**, demonstrating the necessity of the cavitand template in quartet assembly and rules out intramolecular NOE between the enol tautomer of uracil to H5 in **1**. Also, no crosspeak was detected between the *N*-Me and H5 protons of **17** as expected, showing the essential nature of the imino protons and thus hydrogen bonding in the assembly.

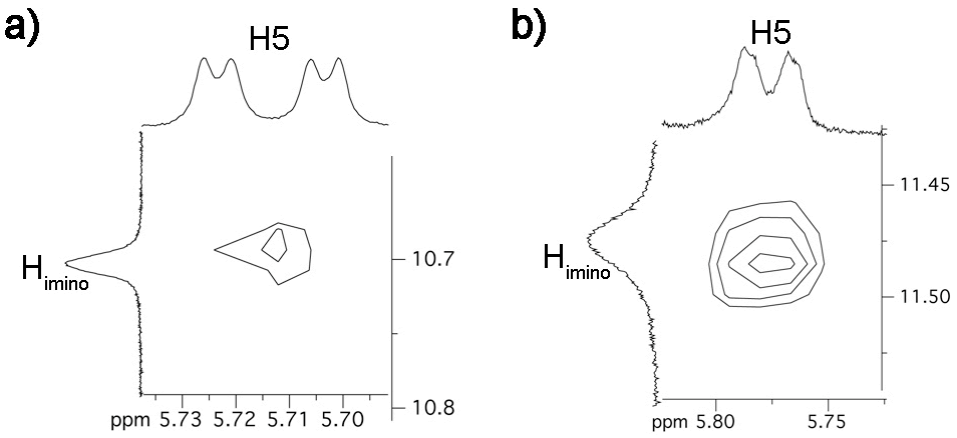


Figure 2.9 Portions of 400 MHz NOE spectrum of 1 in $CDCl_3$ at a) 25 ºC and b) –20 ºC showing crosspeaks between H_{imino} and H5.

Interestingly, the presence of a crosspeak between H6 and H_1' (Figure 2.10a) indicates *syn* glycosidic bonds (Figure 2.10b). This is unexpected as such a configuration usually places the nucleobase and ribose moieties in direct steric conflict. In the case of uracil, the O2 atom would extend directly over the surface of the ribose ring. However, CPK modeling demonstrated

the feasibility of the *syn* configuration by accommodating O2 in a proton 'cleft' mapped out by H_2', H_3' and H_5' (Figure 2.10c). This positions H6 and H_1' in a co-planar fashion well within NOE range.

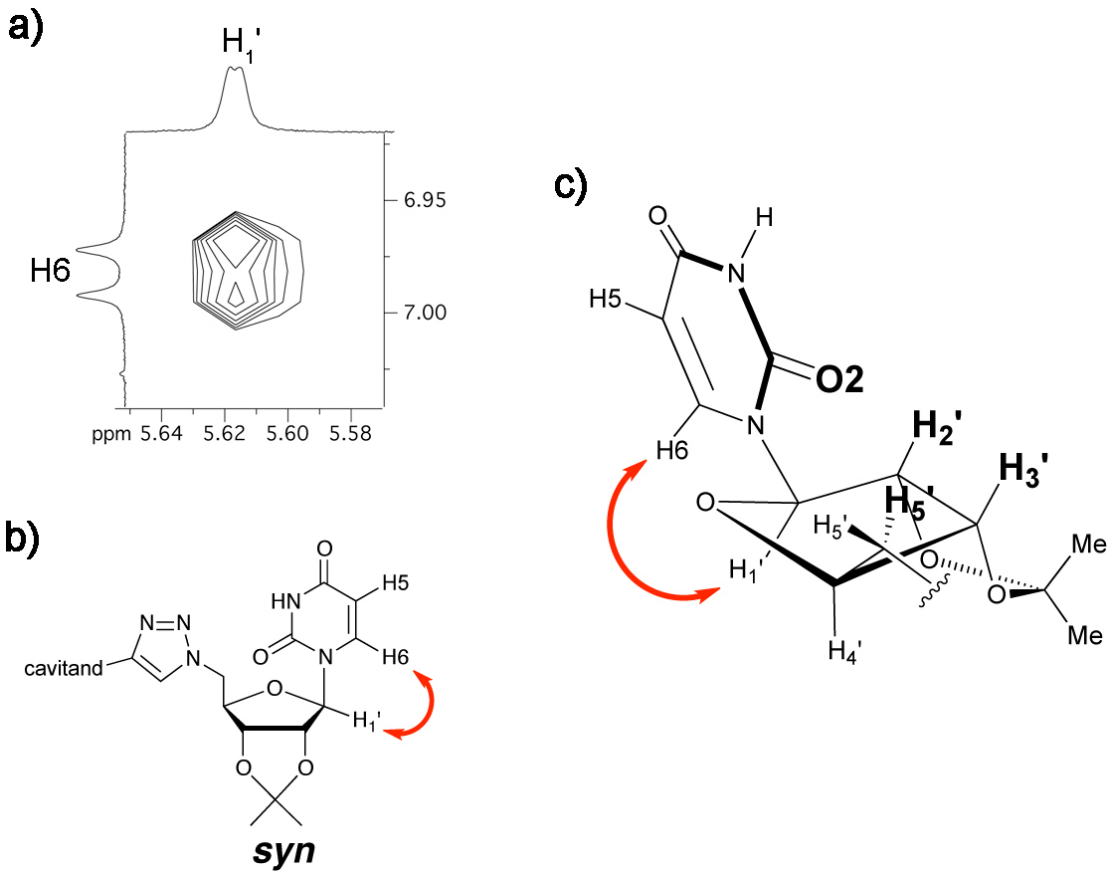


Figure 2.10 a) Portion of 400 MHz NOE spectrum of 1 in $CDCl_3$ at 25 ºC showing crosspeak between H6 and H_1' indicative of *syn* geometry. b) NOE connection shown between H6 and H_1'. c) Model of a uridine residue in 1 showing fitting of the O2 atom in the 'cleft' traced by H_2', H_3' and H_5' protons. H6 and H_1' are coplanar.

As discussed in Section 2.2.1, stacking of the uracil and triazole rings is possible and deemed feasible by CPK modeling. These interactions may confer added stability to the U-quartet. Evidence for this stacking was provided by an NOE crosspeak observed between H6 and

H_c, positioning the two rings in close proximity (Figure 2.11a). No crosspeak is present between H5 and H_c due to their increased spatial distance exceeding the NOE threshold (Figure 2.11b).

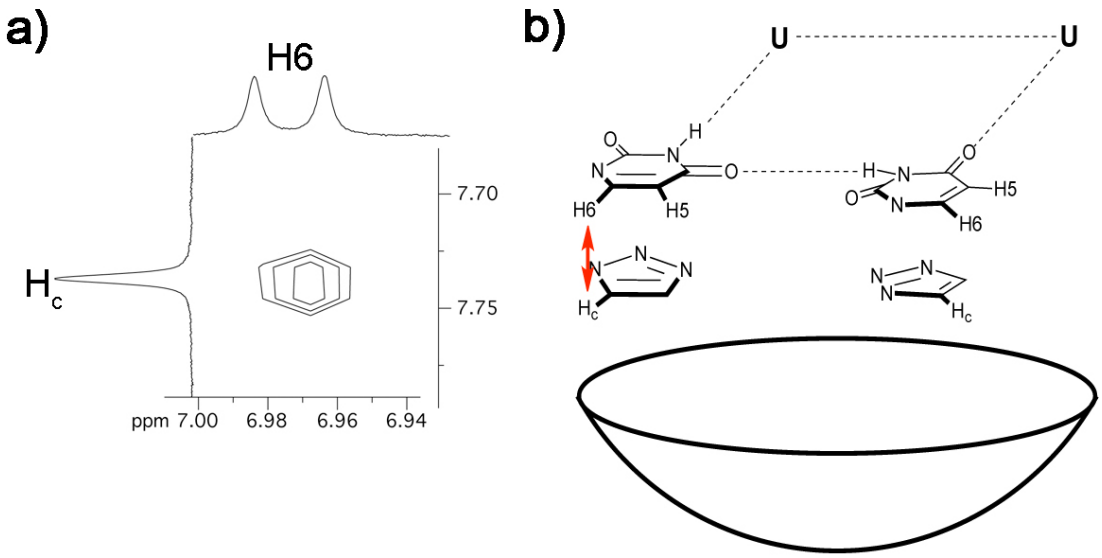


Figure 2.11 a) Portion of 400 MHz NOE spectrum of 1 in $CDCl_3$ at 25 ºC showing crosspeak between H6 and H_c indicative of ring stacking. b) Illustration of proposed U-quartet in 1 with uracil-triazole stacking. NOE between H6 and H_c is shown. Two appendages are displayed. Linkers are omitted for clarity.

2.4.4 Variable-temperature ^{1}H NMR studies

The thermal stability of the U-quartet was assessed by measuring its ^{1}H NMR spectrum at temperatures ranging from –20 ºC to 55 ºC (Figure 2.12). As the temperature is increased, the most pronounced trend is the upfield shift of H_{imino} to about 10.2 ppm at 55 ºC, reflecting its increasingly shielded state as a result of weakening hydrogen bonds not necessarily amounting to complete denaturation.

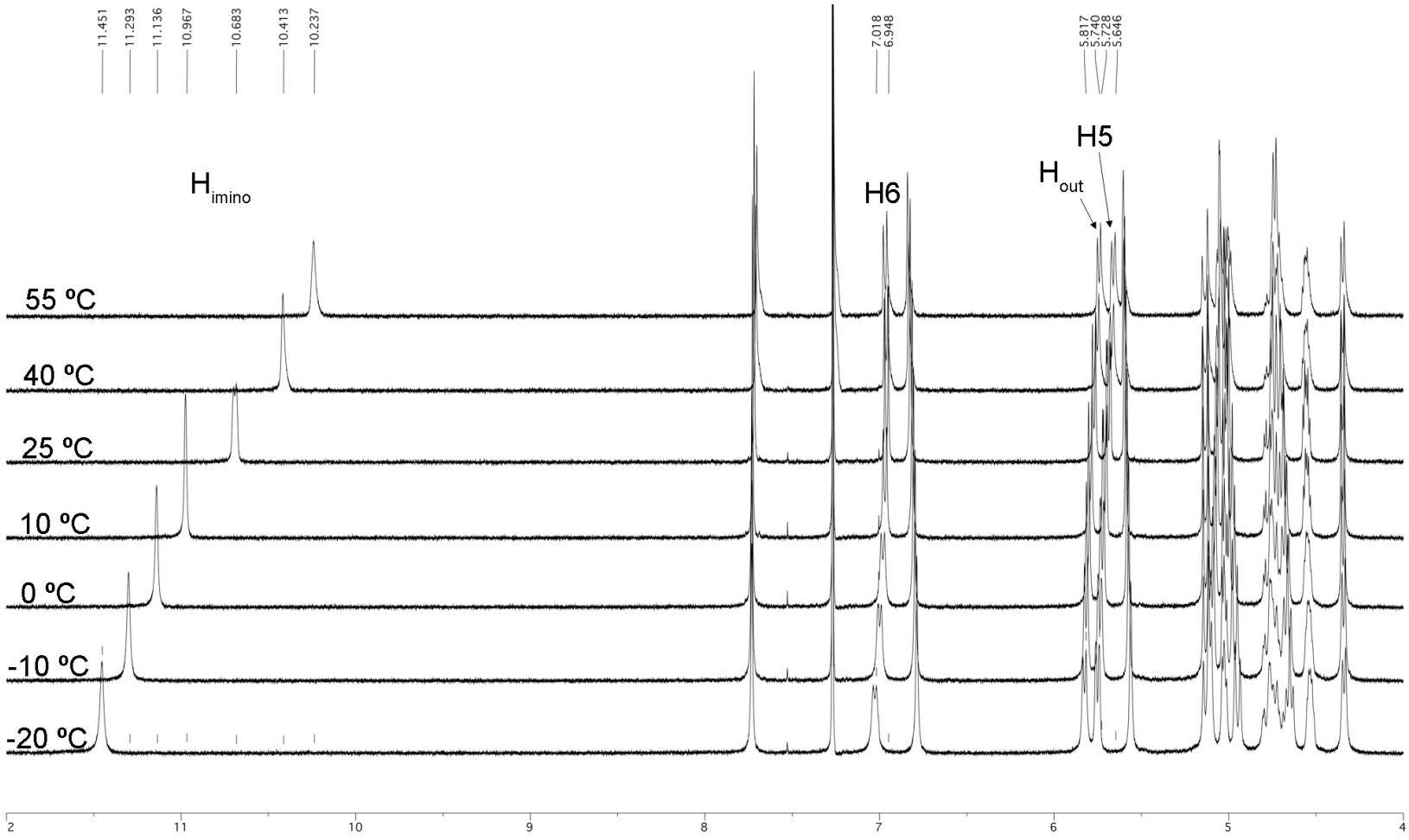


Figure 2.12 Variable temperature ^{1}H spectra of 1 in $CDCl_3$.

Small upfield chemical shifts are also observed for several non-exchangeable protons with increasing temperature (H_{out}: $\Delta\delta_H$ = 0.08 ppm, H5: $\Delta\delta_H$ = 0.08 ppm, H6: $\Delta\delta_H$ = 0.07 ppm). These are likely caused by conformational changes arising from hydrogen bond weakening, positioning these protons further within the anisotropic field of the cavitand ring. Nevertheless, the VT spectra demonstrate that some of the integrity of the U-quartet is maintained at elevated temperatures.

2.4.5 Diffusion NMR studies

Diffusion is the translational motion of particles through a liquid or gas. As part of independent studies on Brownian motion, Einstein and Smoluchowski first described the phenomenon as a function of temperature, now known as the Einstein relation (Equation 1).

$$D = \mu K_B T \tag{1}$$

Where D and μ denote the diffusion constant and mobility of the particle respectively. The Stokes-Einstein equation is a variant of the Einstein relation applicable to the diffusion of spherical particles in a liquid of low Reynolds number (Equation 2).

$$D = \frac{K_B T}{6\pi\eta r} \tag{2}$$

Where η represents the viscosity of the liquid and r is the hydrodynamic (or Stokes) radius of the particle. This equation therefore provides a good approximation of the Van der Waal's radius of a spherical particle if its diffusion constant D is known. As D is also proportional to molecular weight as determined by Chapman et al. (Equation 3),[107]

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

the oligomeric state of the species can be determined by comparing its *D* value to that of a control compound, usually one that is structurally similar but incapable of intermolecular aggregation due to the lack of required functionalities.

Several methods exist for the experimental determination of *D*. Among them, diffusion-ordered NMR spectroscopy (DOSY) has distinguished itself in recent years due to its practical convenience and expeditious delivery of highly accurate measurements as a result of advances in NMR technology. The technique allows molecular diffusion to be quantified by the application of magnetic pulsed field gradients along the length of the sample.[108] Most modern spectrometers are fitted with variable *z*-gradient probes to provide this capability. The basic gradient spin echo pulse (SE) sequence developed by Stejskal and Tanner applies a field gradient after the initial 90º*x* radiofrequency pulse. The gradient causes a dephasing of magnetization. After a certain diffusion time Δ, a second field gradient is applied to refocus the magnetization (Figure 2.13). The physical translation of nuclei during Δ results in the attenuation of signal intensity,[109] the degree of which depends on the durations of Δ and the gradient pulse (δ), as well as the gradient strength (*g*) as defined by the relation[110]

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\frac{\Delta - \delta}{3})} \tag{4}$$

where *I* denotes the observed signal intensity, I_0 the original signal intensity, *D* the diffusion constant and γ the gyromagnetic ratio of the observed nucleus. Other pulse sequences have since been developed and are routinely used, such as the bipolar pulse longitudinal eddy current delay

(BPLED) sequence[110] that compensates for physical phenomena such as eddy currents to improve data accuracy.

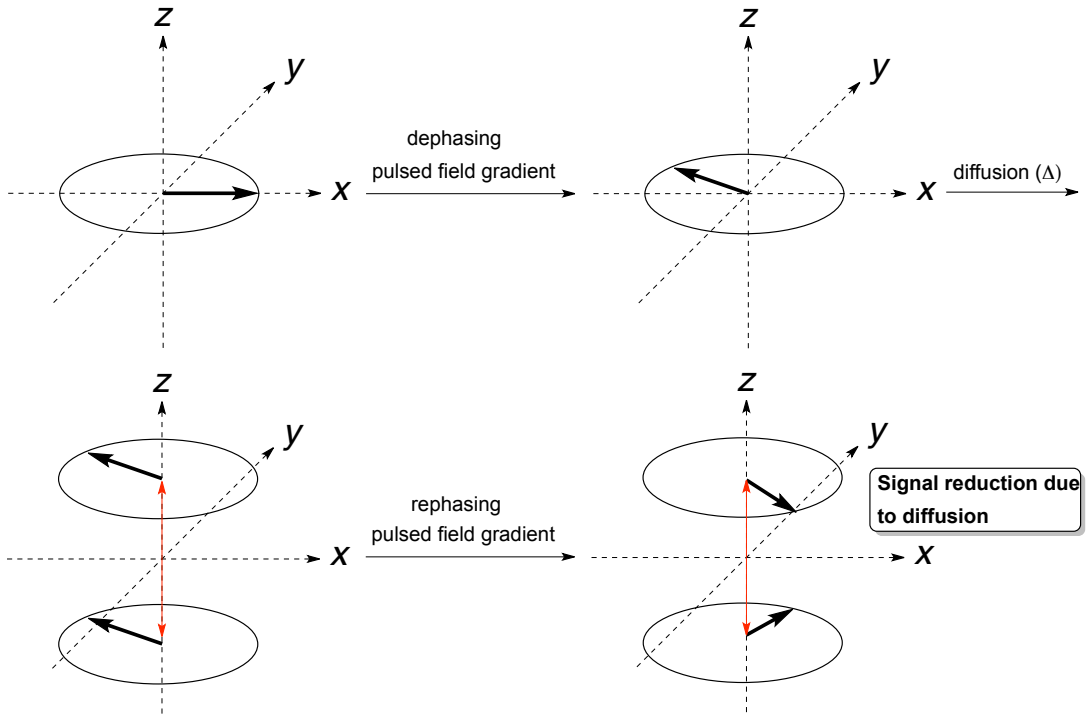


Figure 2.13 Schematic of a basic DOSY spin echo (SE) pulse sequence. Bulk magnetization is represented by a bolded arrow.

In a typical DOSY experiment, values of Δ and δ are adjusted via a series of 1D experiments to obtain at least 95% signal attenuation. With these optimized values, the 2D experiment is conducted with a linear gradient ramp, usually with 16 steps. This allows the fitting of the exponential curve of signal intensity versus gradient strength according to Equation 4. Figure 2.14 shows the curve fit of the H_1' proton of **1** as an example. All DOSY measurements reported in this thesis are extracted from similar curves of selected protons for the corresponding

compounds. The diffusion constant *D* can then be extracted from the slope of the line plotted from Equation 5, derived from Equation 4.

$$\ln \frac{I}{I_0} = -D(2\pi\gamma g\delta)^2(\Delta - \frac{\delta}{3}) = -DQ \tag{5}$$

More conveniently, the advent of software-based algorithms such as SimFit allows direct calculation of *D* values based on longitudinal (T_1) and transverse (T_2) relaxation times.

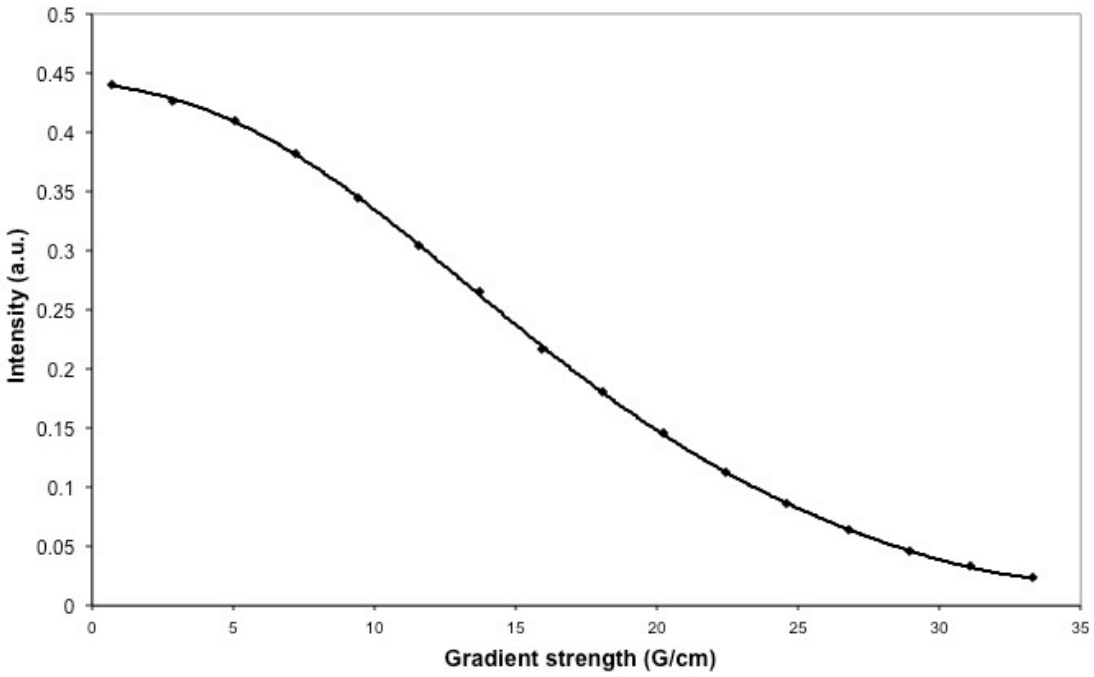


Figure 2.14 Signal intensity decay curve over 16 points for H_1' proton of 1.

As a diagnostic tool for aggregation, DOSY has been used extensively by our group, as well as in the characterization of resorcinarene[111, 112] and rosette-based oligomers[113]. The technique was employed to ascertain the unimolecular nature of **1** in $CDCl_3$ and thus confirm the

existence of a singular, discrete U-quartet in solution. Solutions of **1** and **17** in $CDCl_3$ at identical concentrations (2.7 mM) and temperatures (25 ºC) were analyzed. The H_1' proton of each compound was observed and *D* values were calculated by the SimFit algorithm. Experiments were repeated thrice and the average values with standard error are tabulated in Table 2.3.

Compound	*D* (x10^{-10} m^2 s^{-1})	Ratio
1	3.71 ± 0.10	0.99
17	3.67 ± 0.10	

Table 2.3 H_1' diffusion constants (*D*) of 1 and 17 in $CDCl_3$ at 25 ºC (2.7 mM).

A diffusion constant ratio of close to unity shows that the oligomeric states of the two species are identical. Since **17** is unable to engage in hydrogen bonding, it can be concluded that both compounds are unimolecular in $CDCl_3$ under ambient conditions.

2.4.6 Circular dichroism (CD) studies

Chiral chromophores differentially absorb left and right circularly polarized light as a function of wavelength in a phenomenon known as circular dichroism (CD).[114] The absorption difference can be measured spectroscopically. Asymmetric secondary structures adopted by chiral chromophores exhibit unique spectral profiles enabling characterization of such motifs as α-helices and β-sheets of polypeptides.[115]

Guanine has two absorption bands usually found in the middle-ultraviolet region at 250 and 279 nm attributed to π-π* transitions (2.15a).[116, 117] Exciton coupling due to transition dipole moment alignment of stacked guanine residues in G-quadruplexes cause absorption band splitting into positive and negative CD signals, usually accompanied by signal intensification (Figure 2.15b). Studies on poly(G) systems have observed the positive exciton couplet at 260 nm and the negative at 240 nm.[118-120] These characteristic properties allow identification of these assemblies by CD spectroscopy as an alternative to NMR techniques.

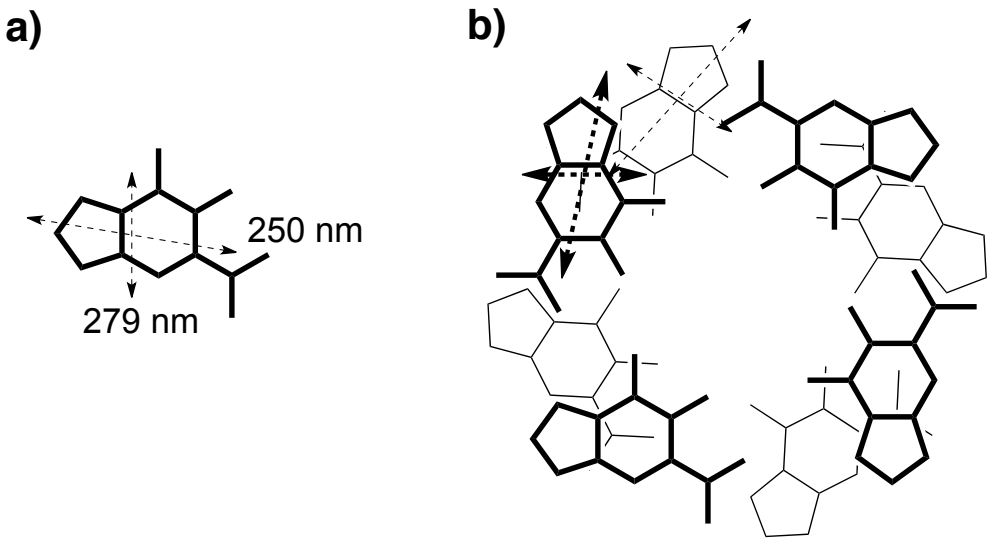


Figure 2.15 a) π–π* transition dipoles of a guanine base. b) Overlapping and exciton coupling of transition dipoles in a G-quadruplex.

In contrast, CD data pertaining to uridine is uncommon, and for U-quartets unknown. Eyring and co-workers provided the first spectral traces of uridine and its derivatives in several solvents.[121] Consistent across all the measured spectra are absorption bands with maxima positioned around 200 and 270 nm. Evidence for π-π* transitions in effect was also presented,

although the contribution of *n*-π* transitions was not ruled out. More recently, an extensive computational analysis supports the operation of π-π* transitions in both absorption bands.[122]

The CD spectrum of **1** in chloroform reveals a single positive absorption peak with λ_{max} ~ 245 nm (Figure 2.16). Due to the absorption limit of chloroform, observation below 230 nm is not possible. The presence of a negative peak to suggest exciton coupling in a quartet thus could not be established. On switching the solvent to methanol, a bathochromic shift to λ_{max} ~ 255 nm occurs along with signal dampening, suggesting quartet denaturation. To rule out solvatochromic shifting, spectra of **17** were obtained in chloroform and methanol and were found to be indistinguishable with λ_{max} ~ 255 nm. It is noteworthy that the spectrum of **1** in methanol is identical to that of **17** in both chloroform and methanol. This indicates the nonexistence of a quartet in a methanolic solution of **1**.

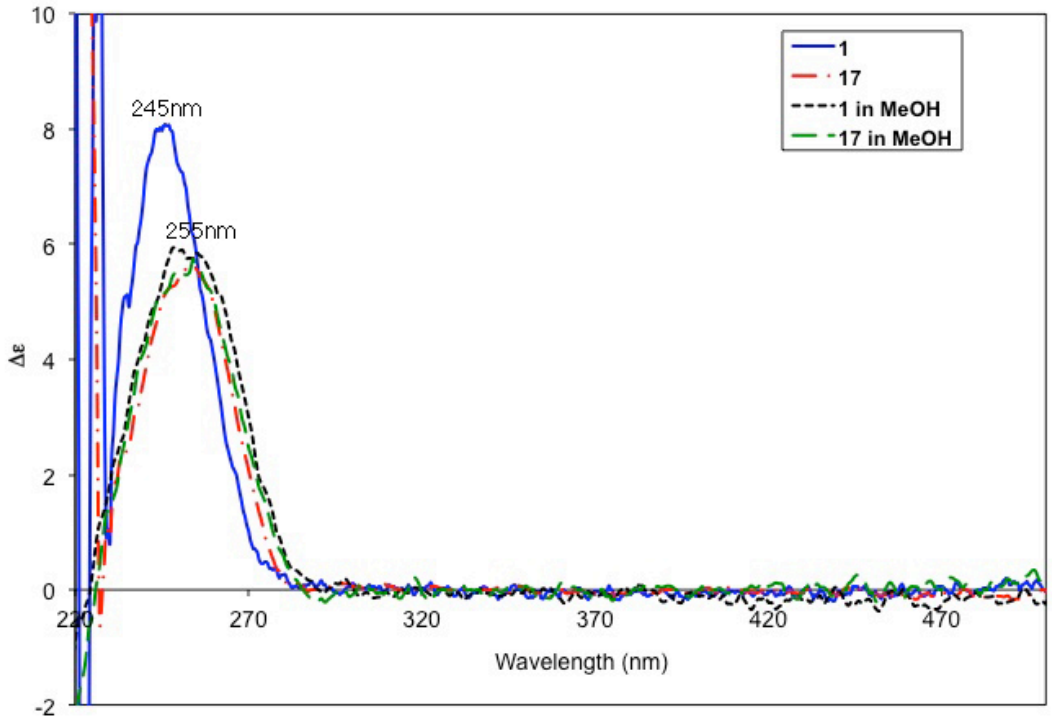


Figure 2.16 CD spectra of 0.1 mM solutions of 1 and 17 in $CHCl_3$ and MeOH.

2.4.6.1 Cation extraction studies

Cation binding has been displayed by several G-quadruplex-supported U-quartets described in Chapter 1 and is a property allowing their potential use as artificial ionophores. The behaviour of **1** in the presence of metal cations was investigated. This was performed by treating chloroform solutions of **1** with the picrate salts of Na^+, K^+ and Sr^{2+} for a duration of one week, after which the solutions were centrifuged, the supernatants collected and diluted to the desired concentration and subjected to CD and NMR spectroscopy. The picrate salts were prepared according to literature procedures.[123]

Na^+ and K^+ were found not to be taken up by **1** i.e. the measured spectra are identical to that of cation-free **1** in chloroform. In the presence of $Sr(pic)_2$ however, an induced CD signal of the picrate anions manifests as a broad absorption band from 300 – 450 nm (Figure 2.17). This suggests the sequestering of Sr^{2+} by the conjugate in chloroform and a conjugate-metal picrate complex (henceforth denoted as $\mathbf{1}_2 \bullet Sr^{2+}$, stoichiometry explained below) assembles, positioning the picrate anions in a chiral environment. λ_{max} for this species was observed at ~ 245 nm, indicating the continued presence of the U-quartet. On switching the solvent to methanol, the picrate signal is extinguished and a bathochromic shift occurs with the new λ_{max} ~ 255 nm, indicating the denaturing of both the complex and U-quartet.

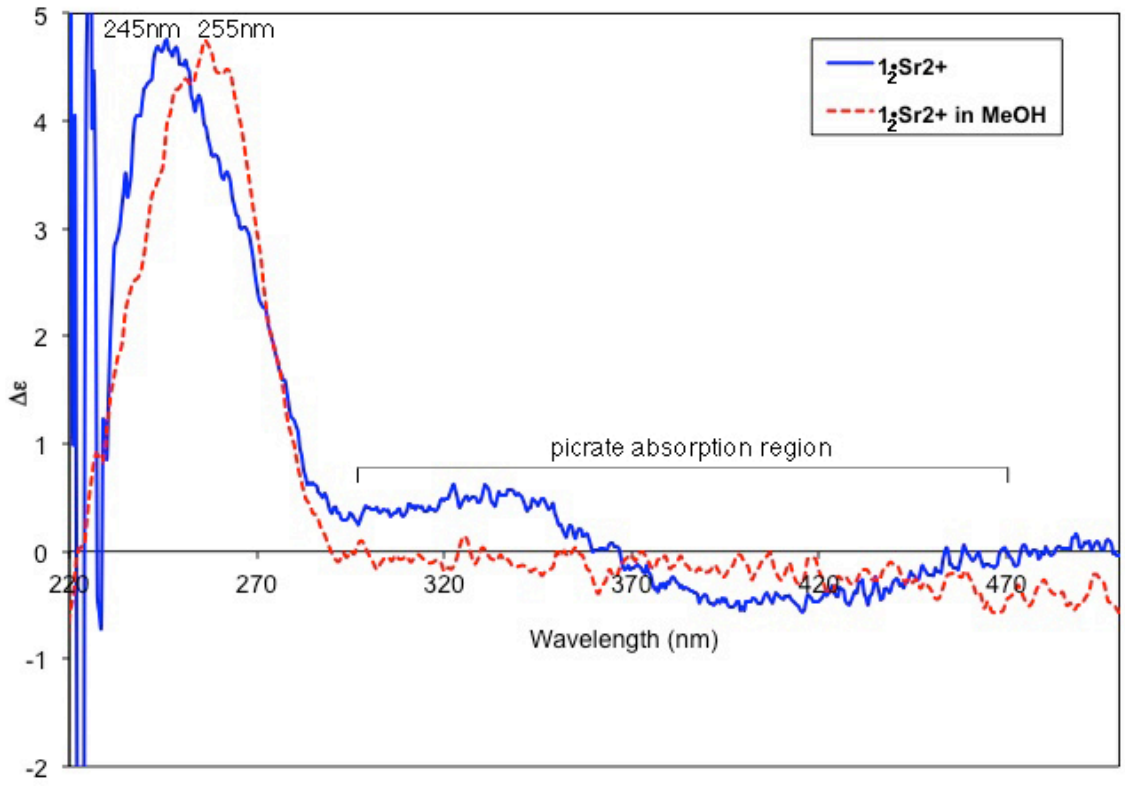


Figure 2.17 CD spectra of 0.1 mM solutions of $\mathbf{1}_2$•Sr^{2+} in $CHCl_3$ and MeOH.

The ^{1}H NMR spectrum of the complex shows the picrate aryl proton signal at ~ 8.7 ppm. Integration of the signal reveals a 2 : 1 stoichiometric ratio of the conjugate to $Sr(pic)_2$ (Figure 2.18), indicating the dimeric nature of $\mathbf{1}_2$•Sr^{2+}. This was confirmed by DOSY measurements, with the observed *D* value of H1' for the complex being 84% that of **1** (Table 2.4). This is in close agreement with the theoretical ratio (80%) derived from Equation 3 (Section 2.4.5) for bimolecular with respect to unimolecular spherical systems.

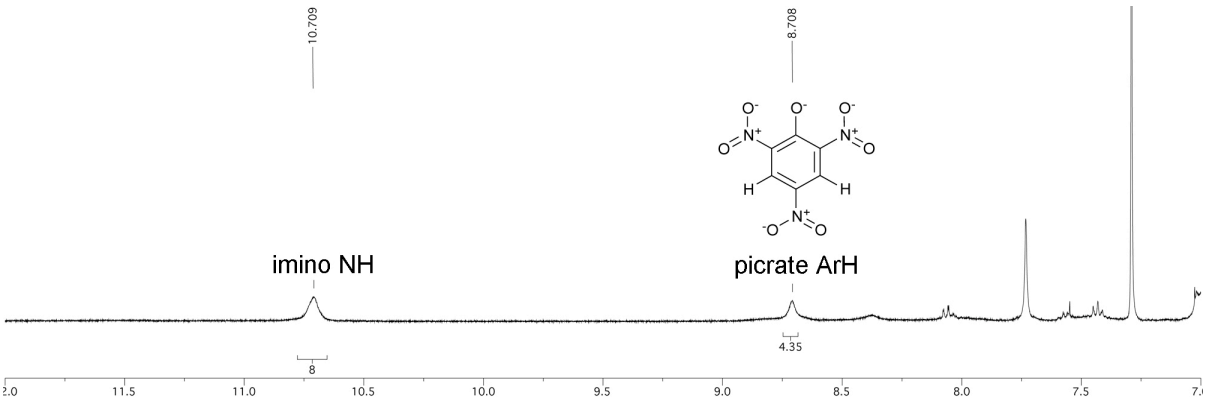


Figure 2.18 Portion of ^{1}H NMR spectrum of $\mathbf{1}_2$•Sr^{2+} showing the imino and picrate aryl proton signals.

Compound	*D* (x10^{-10} m^2 s^{-1})	Ratio
1	3.71 ± 0.10	0.84
$\mathbf{1}_2$•Sr^{2+}	3.11 ± 0.10	

Table 2.4 H_1' diffusion constants (*D*) of 1 and $\mathbf{1}_2$•Sr^{2+} in $CDCl_3$ at 25 ºC (2.7 mM).

Attempts to crystallize the complex from a variety of solvent systems for X-ray crystallographic analysis proved unsuccessful. The complex is predicted to be a symmetric homodimer with two interfaced U-quartets from two molecules of **1** sandwiching a Sr^{2+} cation, with the picrate anions loosely associated with the assembly (Figure 2.19).

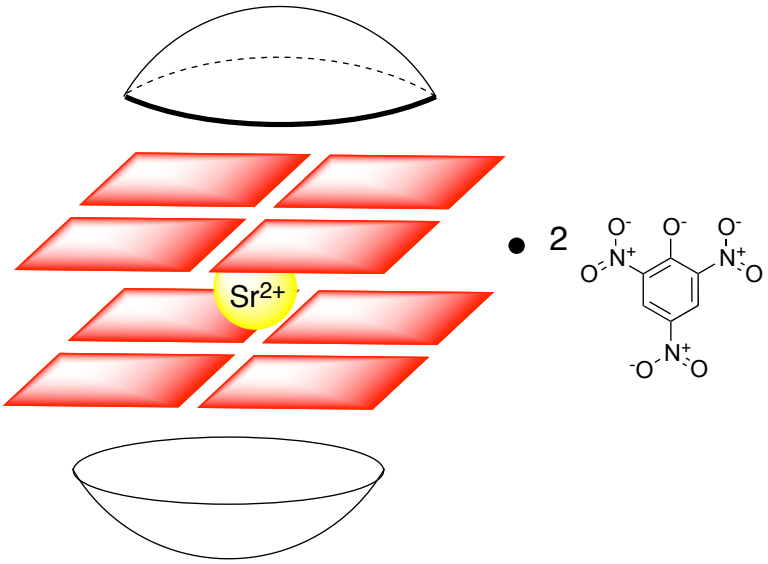


Figure 2.19 Proposed structure of 1_2•Sr^{2+} complex. Uracil residues are represented by red slabs.

2.5 Experimental section

2.5.1 General information

^{1}H NMR spectra were measured on a Bruker Avance 400 MHz spectrometer in $CDCl_3$ [using $CHCl_3$ (for 1H, δ = 7.26) as internal standard] or in DMSO-*d*6 [using DMSO (for 1H, δ = 2.50) as internal standard]. Chemical shifts are reported in ppm from tetramethylsilane. ^{13}C NMR spectra were measured on a Bruker Avance 400 MHz spectrometer in $CDCl_3$ [using $CDCl_3$ (for 13C, δ = 77.0) as internal standard. Chemical shifts are reported in ppm from tetramethylsilane. The following abbreviations are used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. 2D NOESY spectra were acquired with t_{mix} = 800 ms and d_1 = 1500 ms. COSY45 spectra were obtained with d_1 = 1500 ms. DOSY experiments were carried out on a Bruker Avance 400inv spectrometer equipped with a 5mm BBI Z-gradient

probe (inverse broadband probe with z-gradient coil). All measurements were performed using a BPLED gradient pulse sequence (ledbpgp2s). The length of the diffusion gradient was optimized for each sample to obtain at least 95% signal attenuation due to diffusion. Δ and δ values respectively were found to be 65 ms and 4000 ms (for **1** and **1**•Sr^{2+}), 65 ms and 5000 ms (for **17**). Eddy current (t_e) was set at 5 ms. All measurements were taken at 298K with sample concentrations of 2.7 mM. Diffusion coefficients were generated using the SimFit function on Bruker XWinNMR software. MALDI-TOF mass analyses were performed on a Bruker Biflex IV spectrometer in the reflectron mode using 2,5-dihydroxybenzoic acid (DHB) as the matrix. Circular dichroism (CD) spectroscopy was performed on a Jasco J-810 spectrophotometer. Each spectrum is an average of three scans corrected for the baseline. All spectra were acquired with samples in a 1 mm path length quartz cuvette. Flash column chromatography was performed using Silicycle 60 silica gel and eluting solvents were used directly from their commercial bottles. Solvents and reagents for reactions were purchased commercially and used without further purification.

2.5.2 Synthesis of conjugate 1 (similar procedure employed for the synthesis of 17)

To a stirred solution of cavitand **7** (35.6 mg, 0.0345 mmol) and nucleoside **10** (46.9 mg, 0.152 mmol) in argon-purged DMSO (4 mL) was added a solution of copper(II) sulfate pentahydrate (87 μL, 0.04 M in argon-purged Milli-Q water, 0.00345 mmol) followed by a solution of sodium ascorbate (87 μL, 0.4 M in argon-purged Milli-Q water, 0.0345 mmol). The reaction was stirred at 60 °C for 20 h. The solvent was then removed in vacuo and the residue

suspended in water. A few drops of ammonium hydroxide were added to remove the copper catalyst and the mixture was suction filtered. The residue, crude **1**, was washed with deionized water and allowed to air dry, whereupon it was purified via step gradient flash chromatography (100% ethyl acetate – ethyl acetate : methanol 98 : 2 – ethyl acetate : methanol 95 :5) to afford conjugate **1** in 47% yield (42.5 mg, 0.0163 mmol).

Compound 1: White solid; ^{1}H (400 MHz, $CDCl_3$) δ(ppm) 0.92 (12H, t, *J* = 4 Hz, CH_3 feet), 1.29 – 1.38 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 2.20 (8H, m, CH_2 feet), 4.37 (4H, d, *J* = 8 Hz, H_{in}), 4.56 – 4.60 (4H, m, H4′), 4.71 – 4.79 (12H, m, H5′$_a$ and H5′$_b$ and H_d), 5.02 – 5.05 (8H, m, H3′, H_a or H_b?), 5.08 (4H, dd, *J* = 1.6, 8.0 Hz, H2′), 5.16 (4H, d, *J* = 12.0 Hz, H_a or H_b?), 5.62 (4H, d, *J* = 1.6 Hz, H1′), 5.71 (4H, dd, *J* = 4.0, 8.0 Hz, H5), 5.79 (4H, d, *J* = 8.0 Hz, H_{out}), 6.85 (4H, s, ArH), 6.98 (4H, d, *J* = 8.0 Hz, H6), 7.74 (4H, s, H_c), 10.74 (4H, s, imino H); ^{1}H (400 MHz, DMSO-*d*6) δ(ppm) 0.85 (12H, t, *J* = 6.8 Hz, CH_3 feet), 1.24 – 1.30 (84H, m, long chain aliphatic feet and isopropylidene CH_3), 1.45 (12H, s, isopropylidene CH_3), 2.33 (8H, m, CH_2 feet), 4.28 (4H, d, *J* = 7.2 Hz, H_{in}), 4.40 (4H, m, H4′), 4.58 (4H, t, *J* = 8.0 Hz, H_d), 4.65 (4H, dd, *J* = 7.2, 14.0 Hz, H5′$_b$), 4.75 (4H, dd, *J* = 5.2, 14.0 Hz, H5′$_a$), 4.89 (4H, dd, *J* = 4.0, 6.4 Hz, H3′), 4.94 (8H, s, H_a and H_b), 5.12 (4H, dd, *J* = 1.6, 6.4 Hz, H2′), 5.61 (4H, d, *J* = 8.0 Hz, H5), 5.77 (4H, d, *J* = 2.0 Hz, H1′), 5.93 (4H, d, *J* = 7.2 Hz, H_{out}), 7.29 (4H, s, ArH), 7.64 (4H, d, *J* = 8.0 Hz, H6), 8.20 (4H, s, H_c), 11.46 (4H, s, imino H); ^{13}C (100 MHz, $CDCl_3$) δ(ppm) 163.8, 151.0, 148.4, 145.4, 144.4, 143.3, 139.1, 124.8, 115.0, 114.7, 103.4, 99.6, 96.3, 86.5, 84.5, 82.0, 67.6, 52.0, 37.1, 32.0, 30.1, 29.9, 29.7, 29.4, 28.2, 27.3, 25.6, 22.9, 14.1, 1.2; MS(MALDI-TOF): Found: *m/z* 2608.0. Calcd for $C_{136}H_{181}N_{20}O_{32}$: $(M+H)^+$ 2608.0.

Compound 17: Yield: 47%. White solid; ^{1}H (400 MHz, $CDCl_3$) δ(ppm) 0.88 (12H, t, *J* = 7.2 Hz, CH_3 feet), 1.25 – 1.28 (84H, m, long chain aliphatic feet and isopropylidene CH_3), 1.54 (12H, s, isopropylidene CH_3), 2.17 (8H, m, CH_2 feet), 3.30 (12H, s, NMe), 4.32 (4H, d, *J* = 7.2 Hz, H_{in}), 4.46 – 4.50 (4H, m, H4′), 4.65 – 4.70 (8H, m, $H5'_a$ and H_d), 4.77 (4H, dd, *J* = 3.6, 14.0 Hz, $H5'_b$), 4.98 – 5.09 (16H, m, H2′ and H3′ and H_a and H_b), 5.60 (4H, d, *J* = 1.2 Hz, H1′), 5.73 – 5.75 (8H, m, H5 and H_{out}), 6.80 (4H, s, ArH), 7.18 (4H, d, *J* = 8.0 Hz, H6), 7.67 (4H, s, H_c); ^{1}H (400 MHz, DMSO-*d*6) δ(ppm) 0.86 (12H, t, *J* = 6.4 Hz, CH_3 feet), 1.25 – 1.28 (84H, m, long chain aliphatic feet and isopropylidene CH_3), 1.47 (12H, s, isopropylidene CH_3), 2.33 (8H, m, CH_2 feet), 3.15 (12H, s, NMe), 4.28 (4H, d, *J* = 7.2 Hz, H_{in}), 4.40 – 4.44 (4H, m, H4′), 4.58 (4H, t, *J* = 8.0 Hz, H_d), 4.64 (4H, dd, *J* = 7.6, 14.0 Hz, $H5'_b$), 4.76 (4H, dd, *J* = 5.2, 14.0 Hz, $H5'_a$), 4.92 – 4.96 (12H, m, H_a and H_b and H3′), 5.13 (4H, dd, *J* = 1.6, 6.4 Hz, H2′), 5.75 (4H, d, *J* = 8.0 Hz, H5), 5.82 (4H, d, *J* = 1.6 Hz, H1′), 5.91 (4H, d, *J* = 7.2 Hz, H_{out}), 7.30 (4H, s, ArH), 7.70 (4H, d, *J* = 8.0 Hz, H6), 8.20 (4H, s, H_c); ^{13}C (100 MHz, $CDCl_3$) δ(ppm) 162.7, 151.1, 148.2, 145.2, 144.3, 141.4, 139.2, 124.4, 114.9, 114.8, 102.4, 99.7, 97.1, 86.7, 84.7, 82.3, 77.4 (overlapped), 67.4, 52.2, 37.2, 32.2, 30.2, 30.1, 29.95, 29.89, 29.6, 28.2, 27.8, 27.3, 25.4, 22.9, 14.3; MS(MALDI-TOF): Found: *m/z* 2686.0 Calcd for $C_{140}H_{188}N_{20}O_{32}Na$: $(M+Na)^+$ 2686.1.

2.5.3 Supplementary ^{1}H and ^{1}H-^{1}H COSY spectra

Spectra begin on the following page.

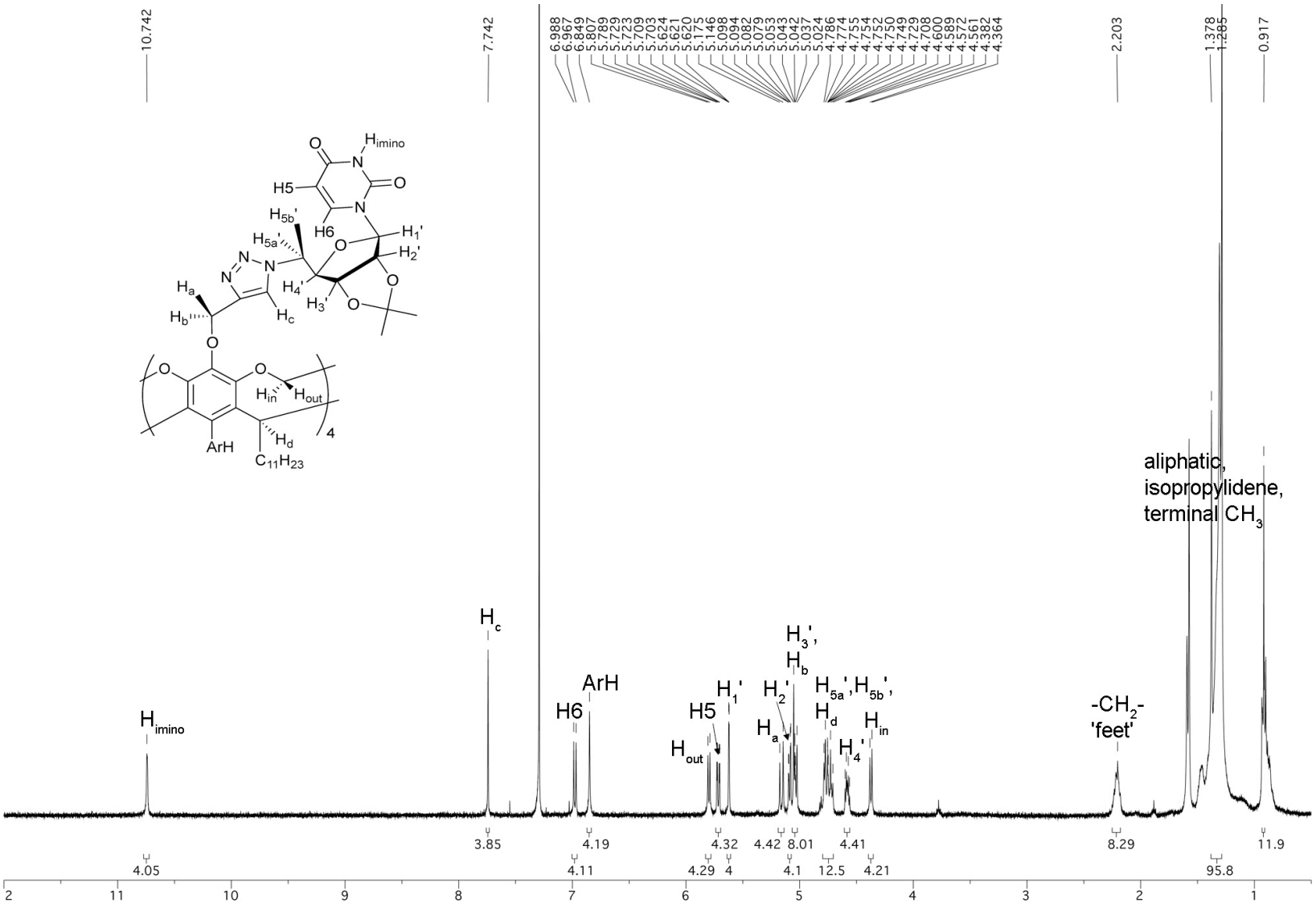


Figure 2.20 ^{1}H NMR spectrum of 1 in $CDCl_3$ at 25 ºC.

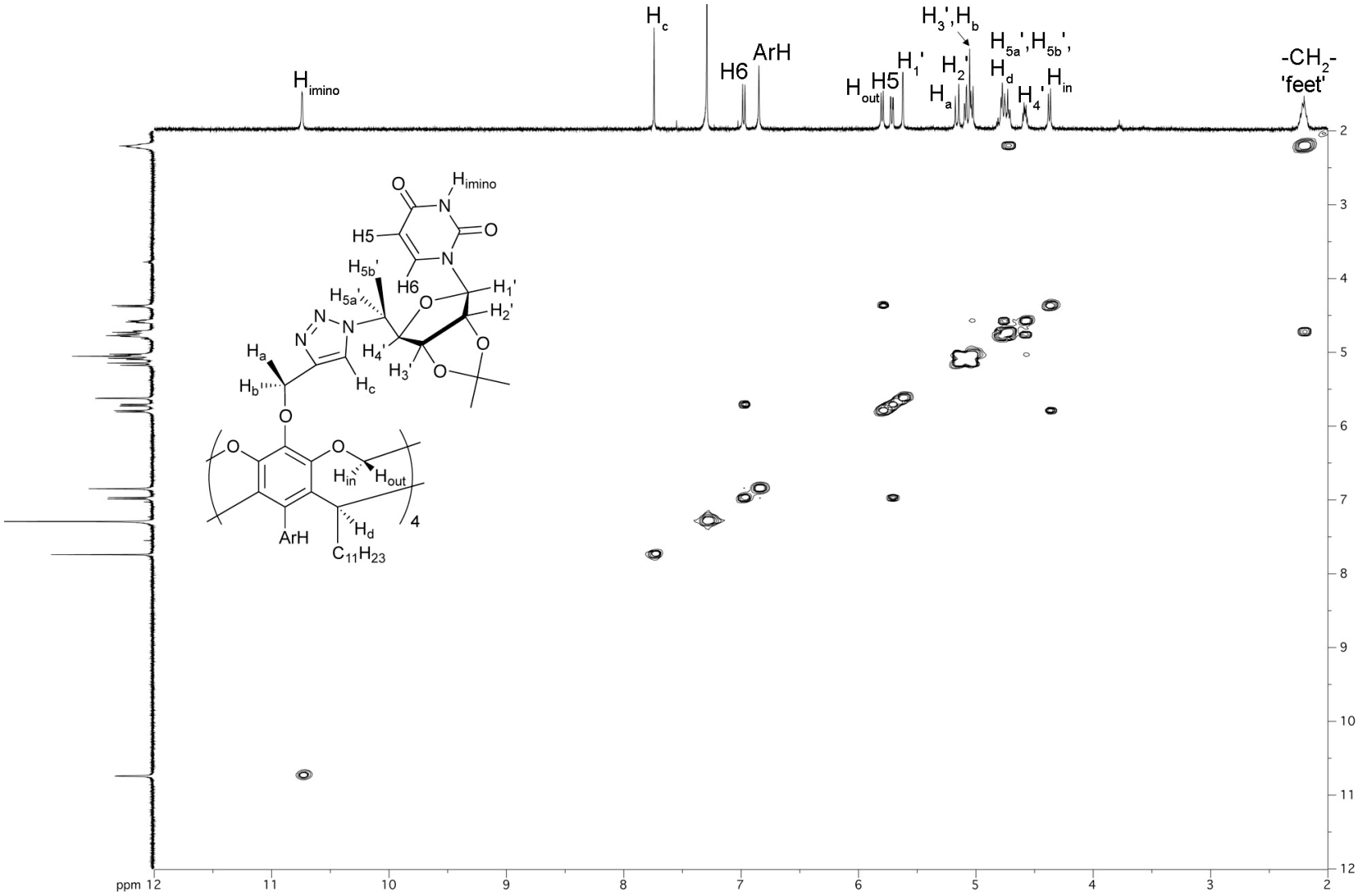


Figure 2.21 ^{1}H-^{1}H COSY spectrum of 1 in $CDCl_3$ at 25 ºC.

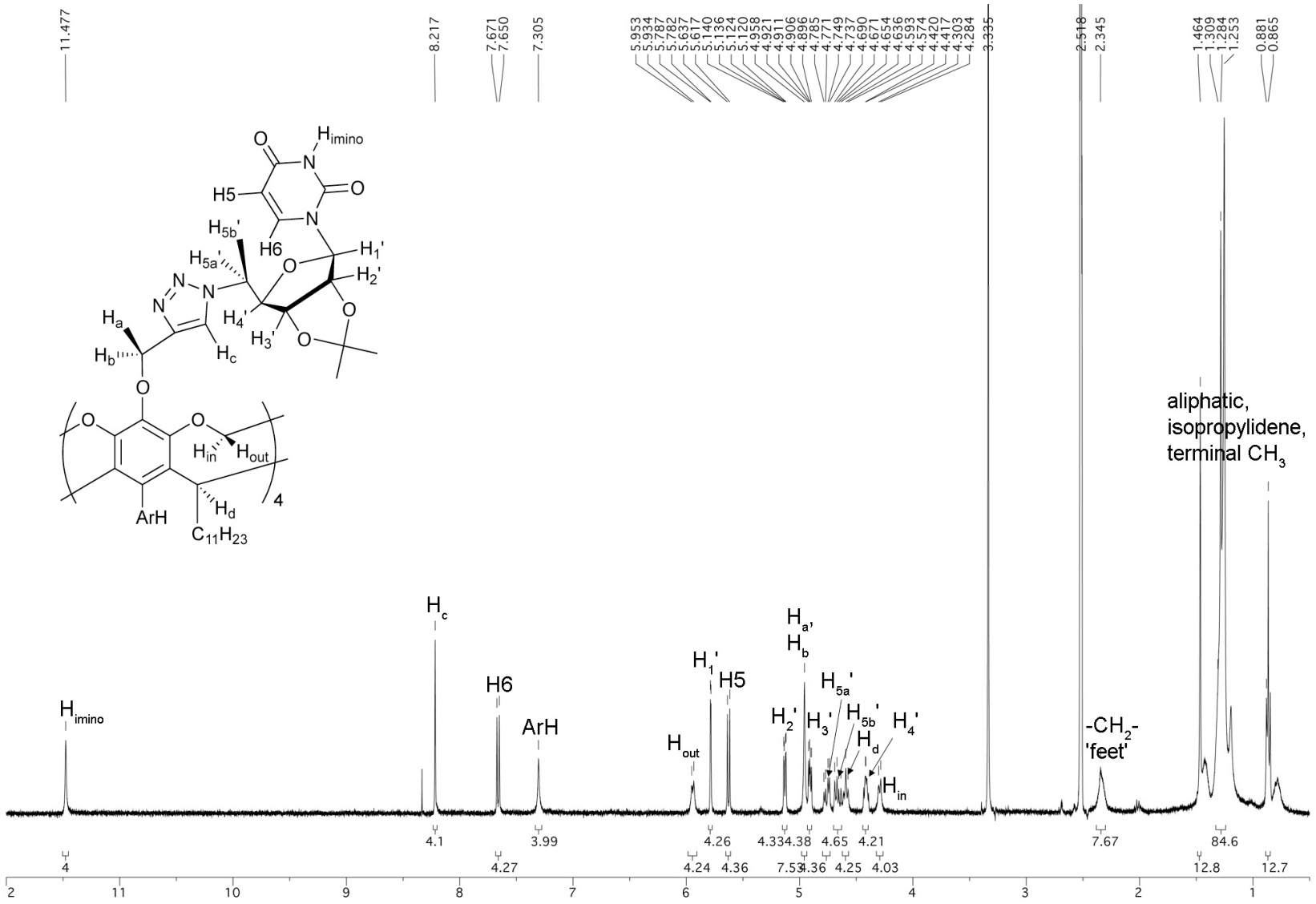


Figure 2.22 ^{1}H NMR spectrum of 1 in DMSO-d_6 at 25 ºC.

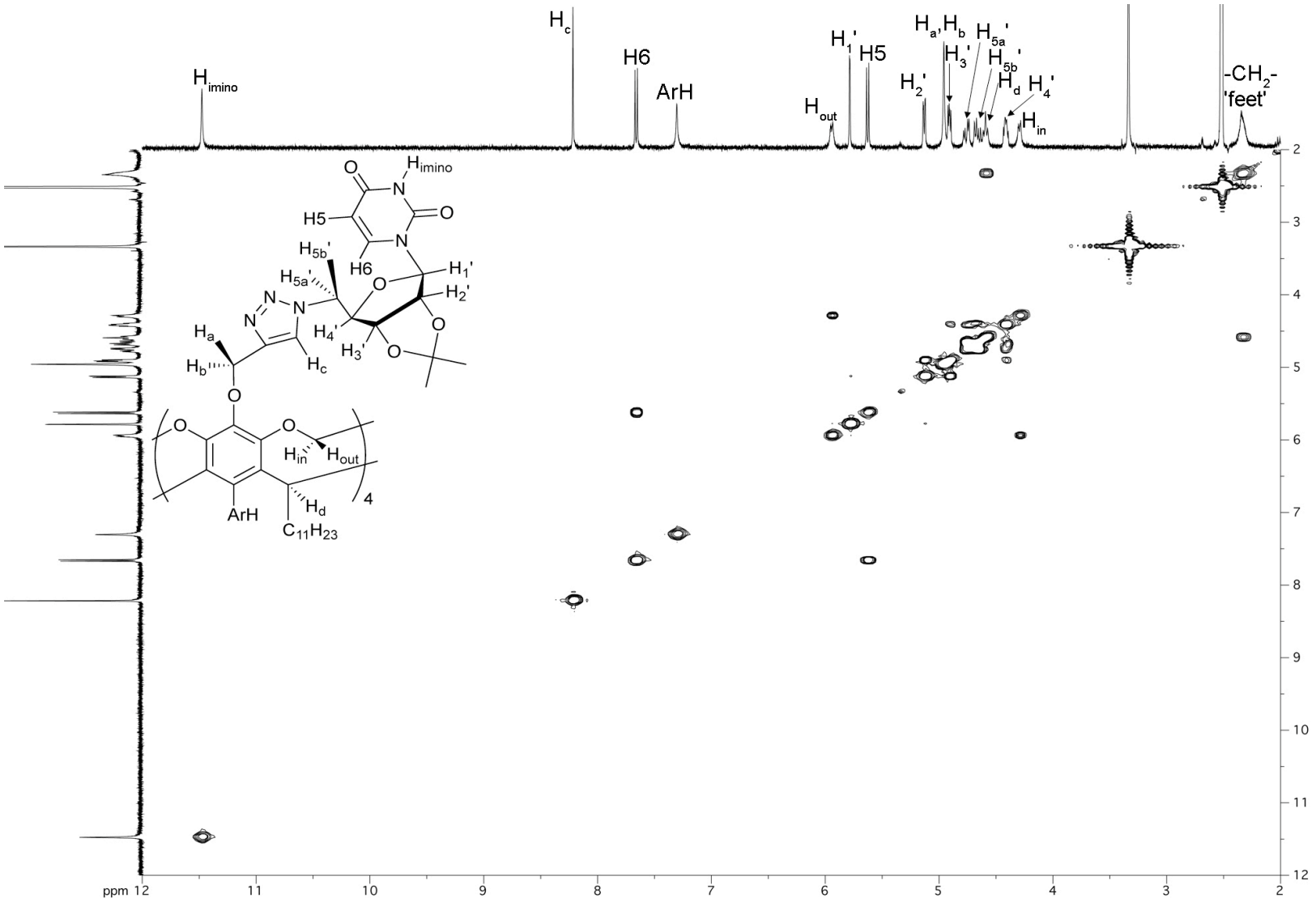


Figure 2.23 ^{1}H-^{1}H COSY spectrum of 1 in DMSO-d_6 at 25 ºC.

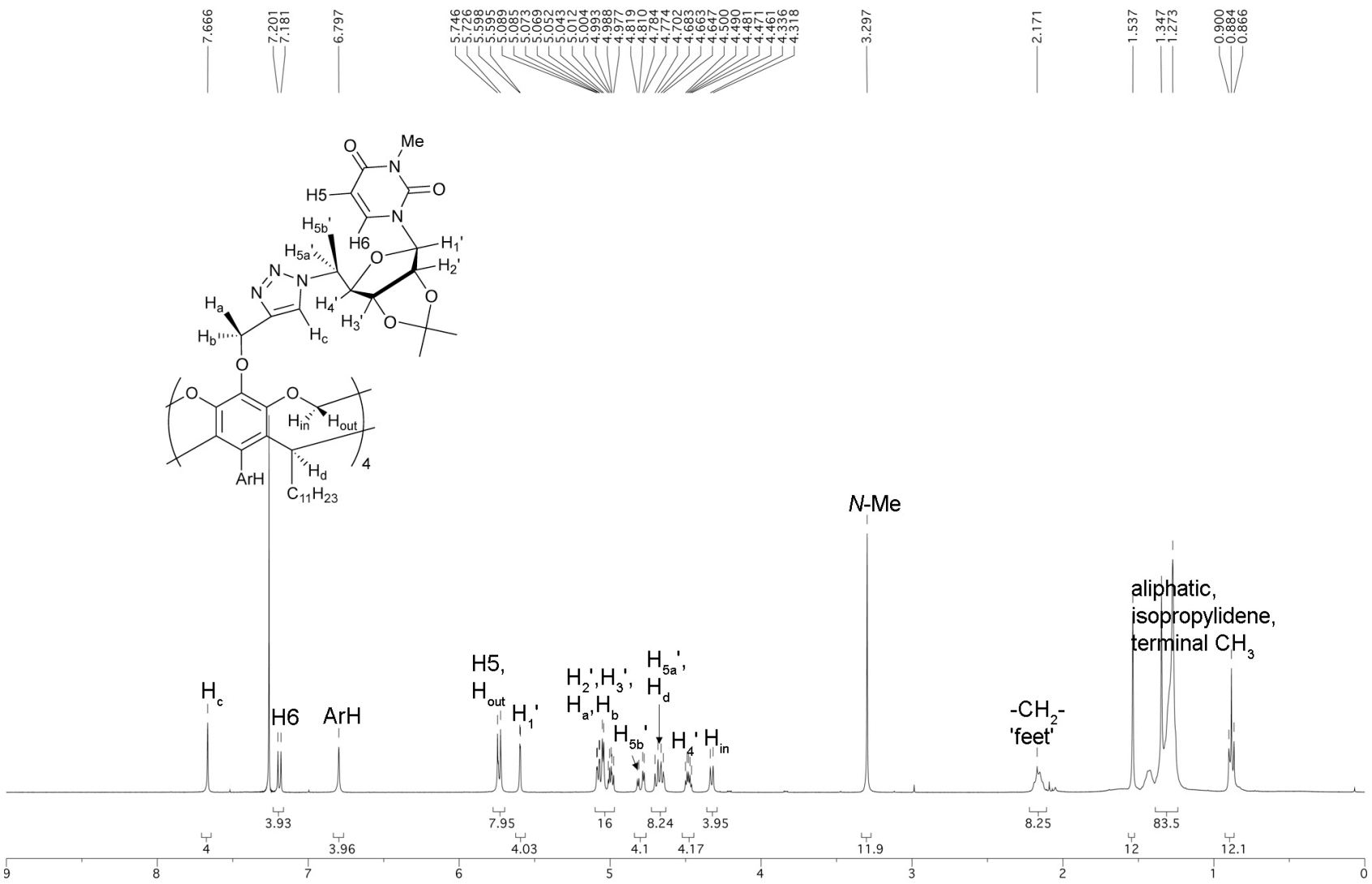


Figure 2.24 ^{1}H NMR spectrum of 17 in $CDCl_3$ at 25 ºC.

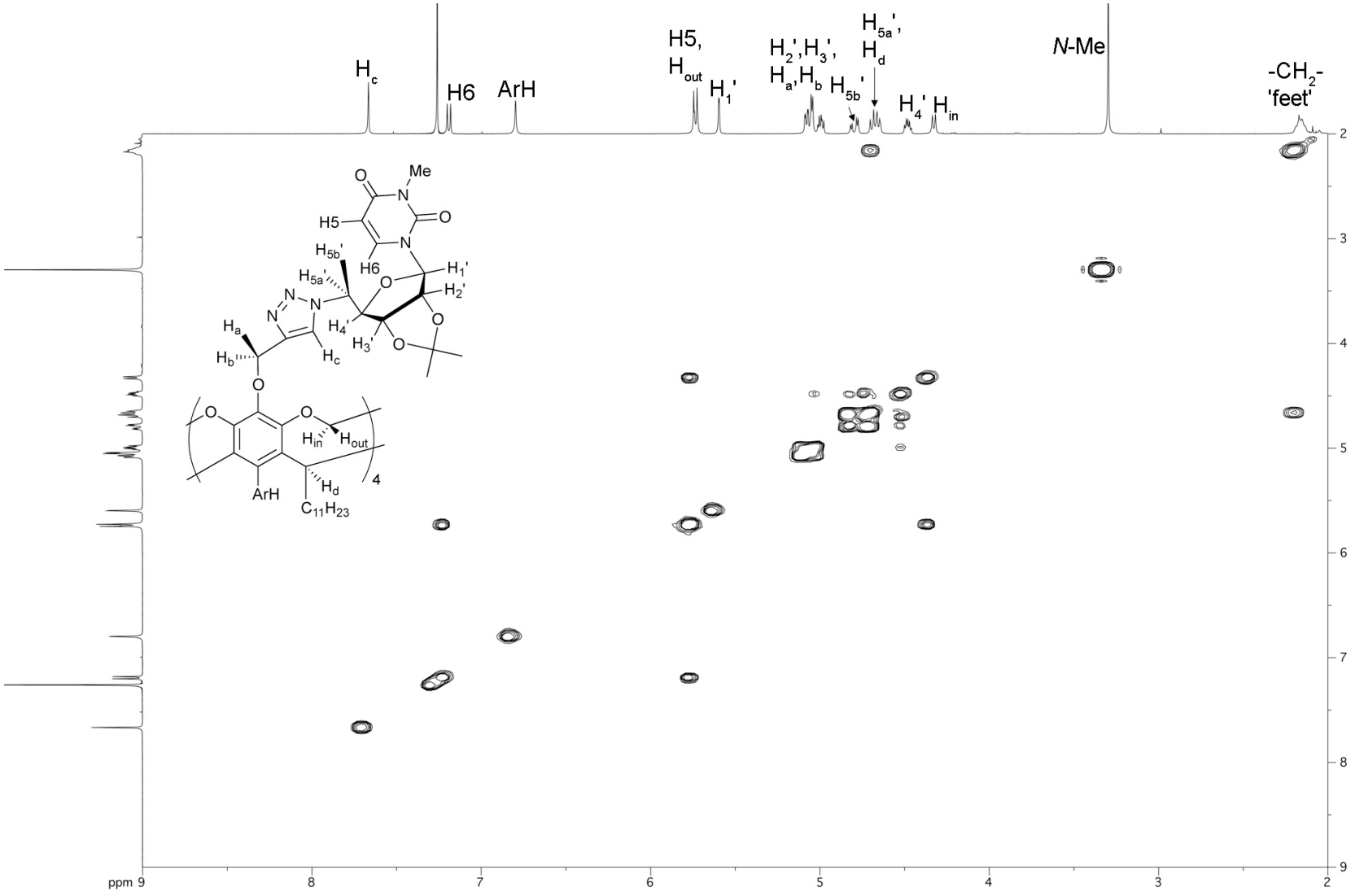


Figure 2.25 ^{1}H-^{1}H COSY spectrum of 17 in $CDCl_3$ at 25 ºC.

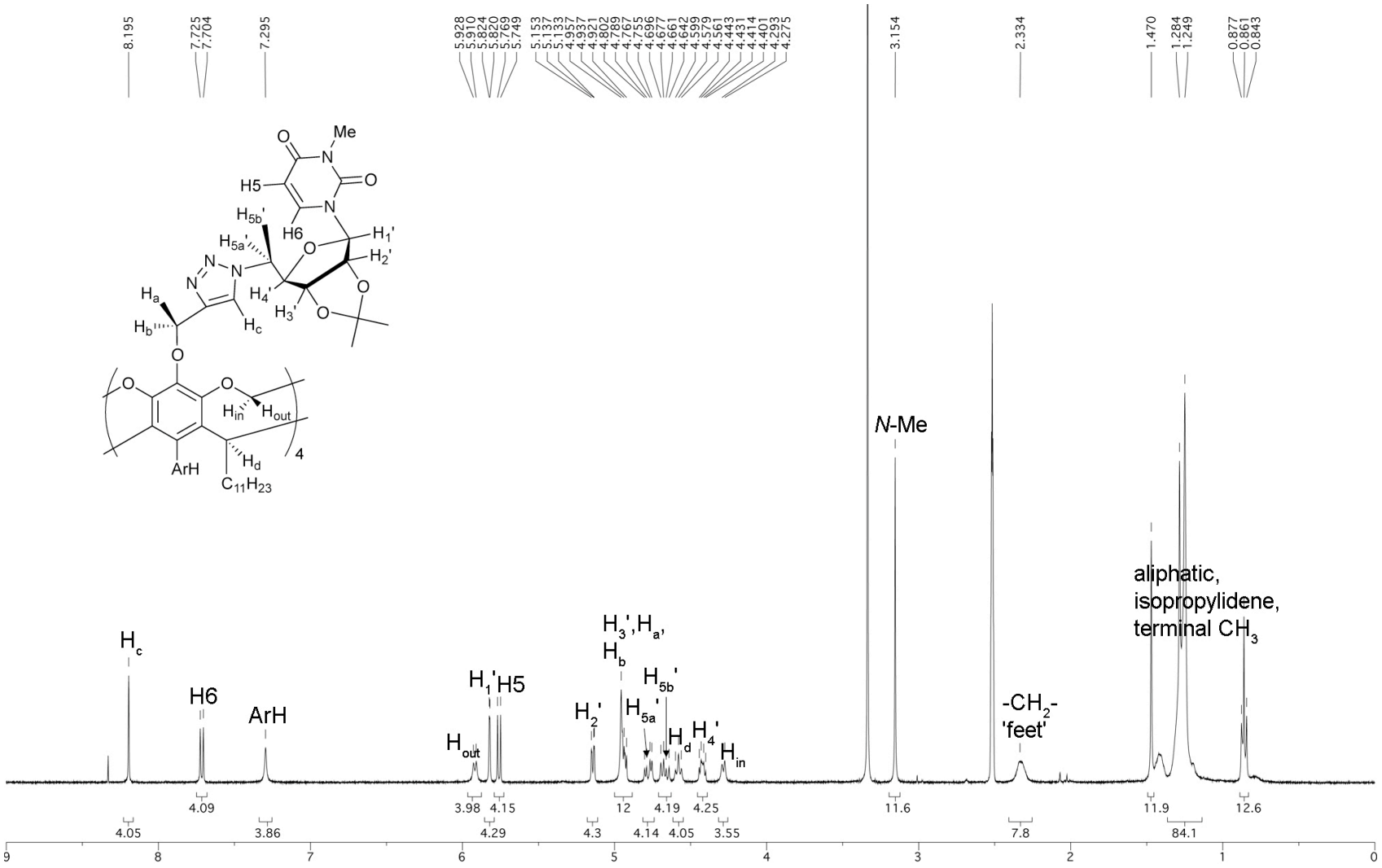


Figure 2.26 ^{1}H NMR spectrum of 17 in DMSO-d_6 at 25 °C.

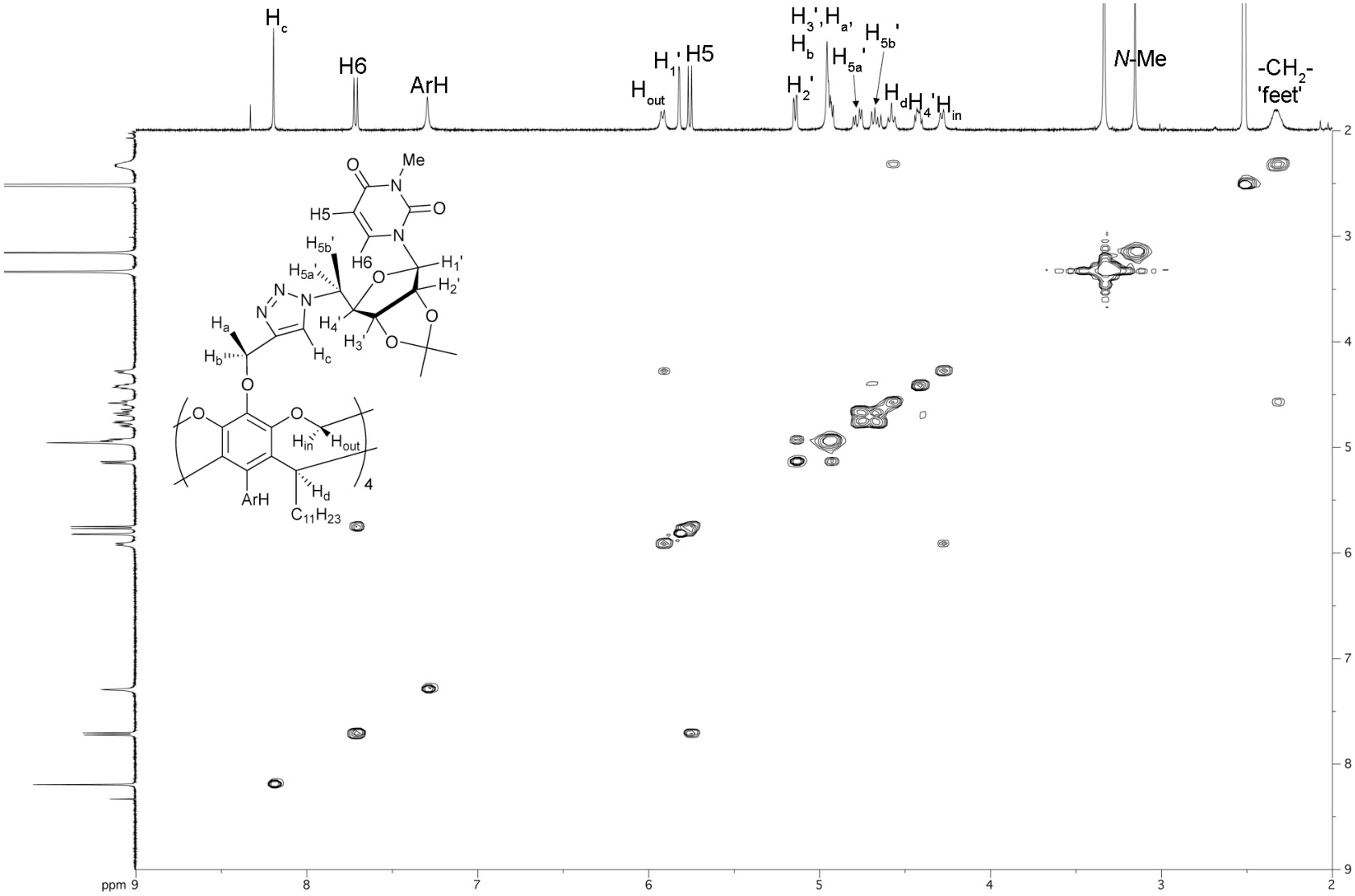


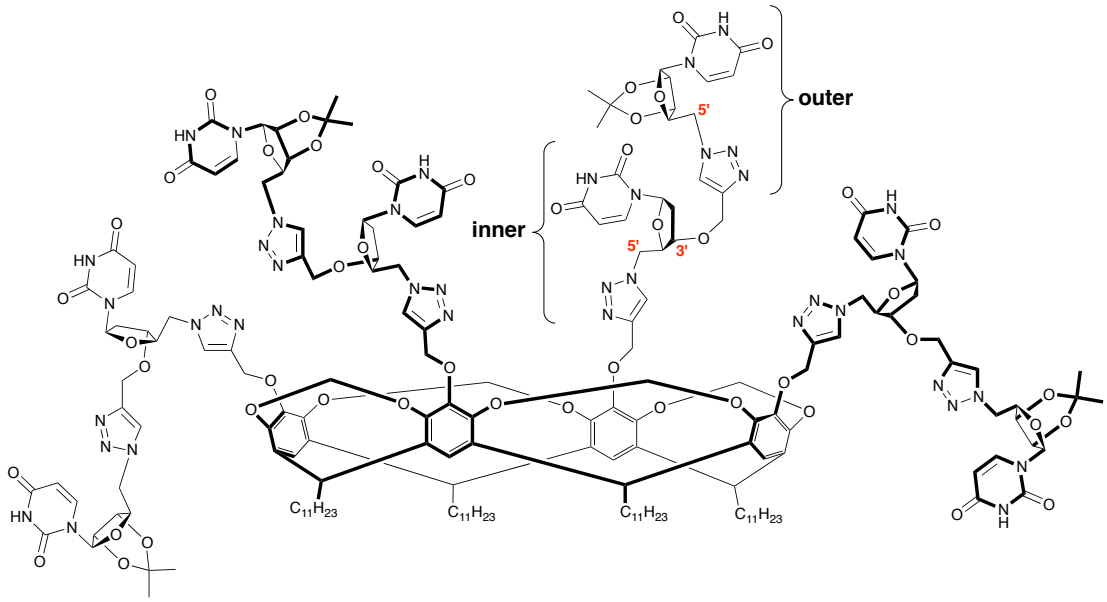
Figure 2.27 ^{1}H-^{1}H COSY spectrum of 17 in DMSO-d_6 at 25 ºC.

Chapter 3: A Template-Assembled Synthetic U-Quadruplex

3.1 Synopsis

In the previous chapter, the ability of a cavitand template to stabilize an inherently unstable U-quartet structure was demonstrated. As an extension of this work, we sought to introduce an additional degree of molecular and structural complexity to the system, not only to pose a synthetic challenge, but also to approach a more natural model on which energetic and synergetic studies might provide more insight on nucleic acid topology and stability. Towards this goal, the construction of a U-quadruplex on our cavitand template was envisaged.

This chapter describes the synthesis and characterization of conjugate **18** (Figure 3.1) in which four uracil-based dinucleoside units are coupled with a lipophilic cavitand. Of note in the design of this molecule is the 5' – 3' directional sense of the dinucleoside moiety, starting from the outer residue. If assembled, an all-parallel two-tiered U-quadruplex would result from this arrangement. As with conjugate **1**, triazole rings provide the backbone of the system. Extensive NMR spectroscopy was performed for complete ^{1}H signal assignment of **18** and to ascertain the assembly of a U-quadruplex in solution.



18

Figure 3.1 Side-on representation of conjugate 18. Outer and inner uridine residues, as well as the 5' and 3' sugar ring positions are labeled to demonstrate directional sense.

3.2 CPK modeling studies

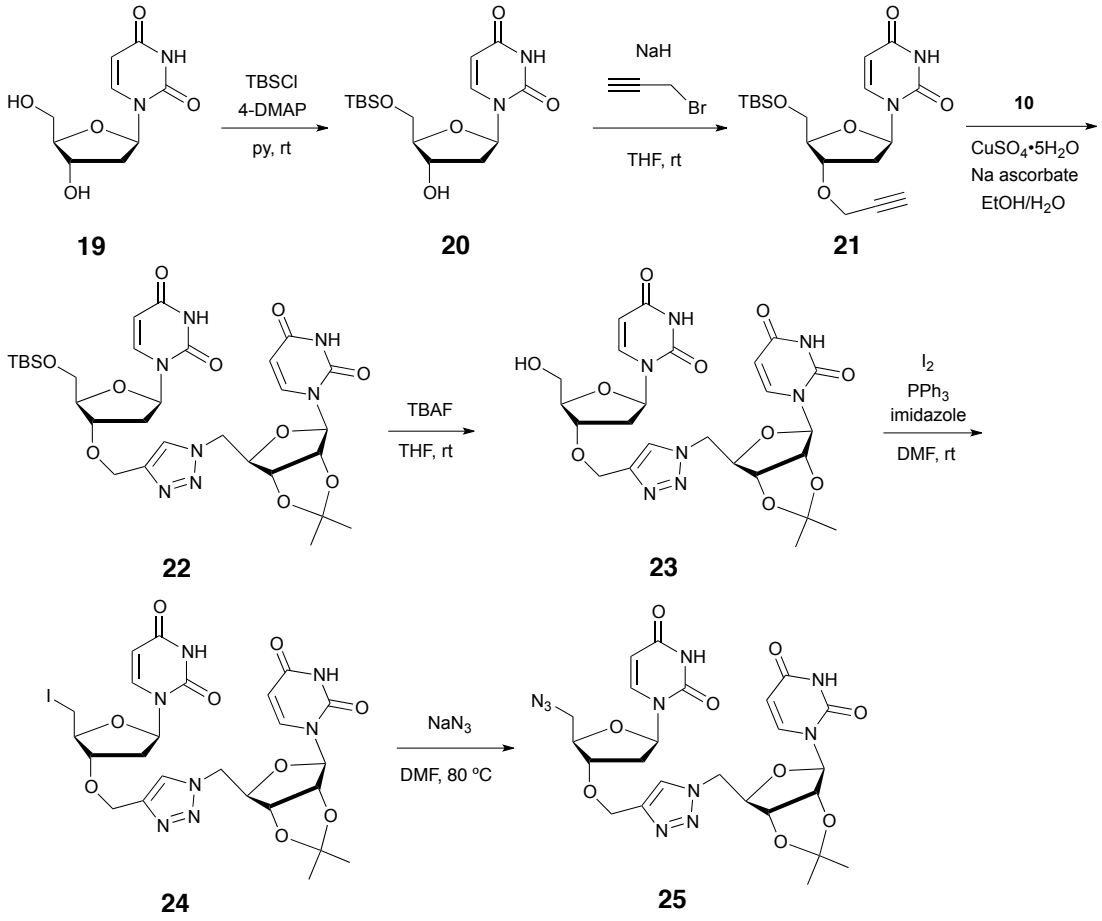
The addition of an outer layer of uridine nucleosides was expected to add a considerable amount of steric bulk and restrict torsional movement of the inner layer, potentially disrupting quartet assembly within and in turn destabilizing a quadruplex assembly. Surprisingly however, CPK modeling of **18** showed that, with an all-*syn* glycosidic bond configuration, a quadruplex could potentially be assembled via the stacking of the two uracil layers with a clearly defined central cavity of size compatible with the common quadruplex-stabilizing cations such as Na^+, K^+ and Sr^{2+}. The model also demonstrated the possibility of triazole ring intercalation to form a fully stacked, alternating triazole-uracil-triazole-uracil scaffold resulting from ring π-π interactions.

3.3 Synthetic strategy

3.3.1 Synthesis of conjugate 18

The preparation of **18** entailed the coupling of cavitand **7** and dinucleoside derivative **25**, itself prepared in several steps (Scheme 3.1). Literature procedures were followed towards 5'-silyl protected intermediate **22**,[124] beginning with the treatment of 2'-deoxyuridine **19** with TBSCl in the presence of DMAP as a nucleophilic catalyst to give 5'-TBS-protected **20**. Generation of the 3'-alkoxide by deprotonation with sodium hydride followed by S_N2 substitution with propargyl bromide provided 3'-propargyl ether derivative **21**, which was then coupled with **10** in a Cu-catalyzed "click" reaction to access **22**.

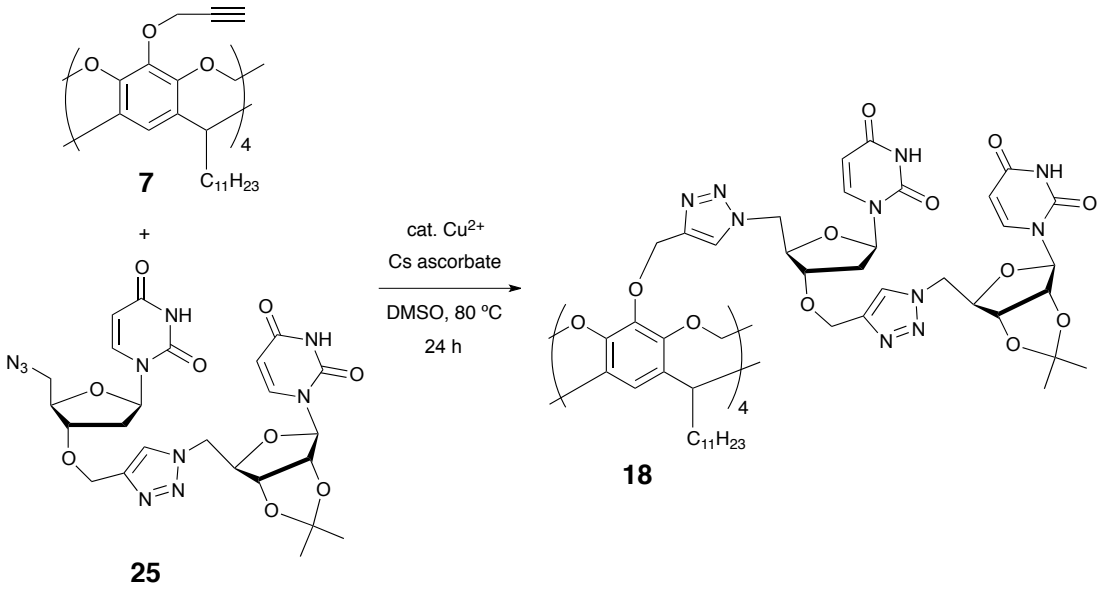
Removal of the 5'-TBS group on **22** was found to proceed cleanly by treatment with TBAF, unmasking the hydroxyl group on **23**. However, direct azidation to **25** via the CBr_4-based Mitsunobu process described for the synthesis of **10** was found to be ineffective here. TLC analysis of the reaction mixture after 15 hours revealed a significant amount of starting material **23** remaining. The sluggish reaction presumably results from the additional uridine unit, which could swivel around to obstruct reagent access to the 5'-OH. Elevating the reaction temperature to 80 ºC could not circumvent this problem and produced unidentified products as detected by TLC. Therefore, an indirect approach was taken in which the Appel iodination of **23** was first performed with less sterically bulky reagents (I_2, imidazole) to access the 5'-I intermediate **24** expediently and cleanly.[125] Subsequent S_N2 substitution with NaN_3 then afforded 5'-azido dinucleoside **25**.



Scheme 3.1 Synthetic route towards dinucleoside 25.

The Cu-catalyzed "click" reaction between **7** and **25** under the same reaction conditions used for the synthesis of **1** was unable to proceed to completion. After 24 hours, TLC analysis revealed multiple spots that were confirmed by MALDI-TOF mass spectrometry to correspond to incomplete conjugation of varying degrees. This is again attributed to the size of **25**, which on successive attachment to **7** would cause increasing steric crowding on the cavitand rim, hindering the approach of subsequent residues of **25** to the propargyl sites (Figure 3.3). Apart from an increase of the reaction temperature, a switch of the reductant from sodium to the more potent cesium ascorbate was made in the hope of driving the reaction to completion. Indeed,

under these conditions the reaction proceeded cleanly after a period of 24 hours to afford the tetra-coupled conjugate **18** as the sole product, confirmed by MALDI-TOF analysis (Scheme 3.2). The compound was found to be soluble in chloroform and only slightly so in methanol.



Scheme 3.2 Synthesis of 18 by the Cu-catalyzed "click" reaction.

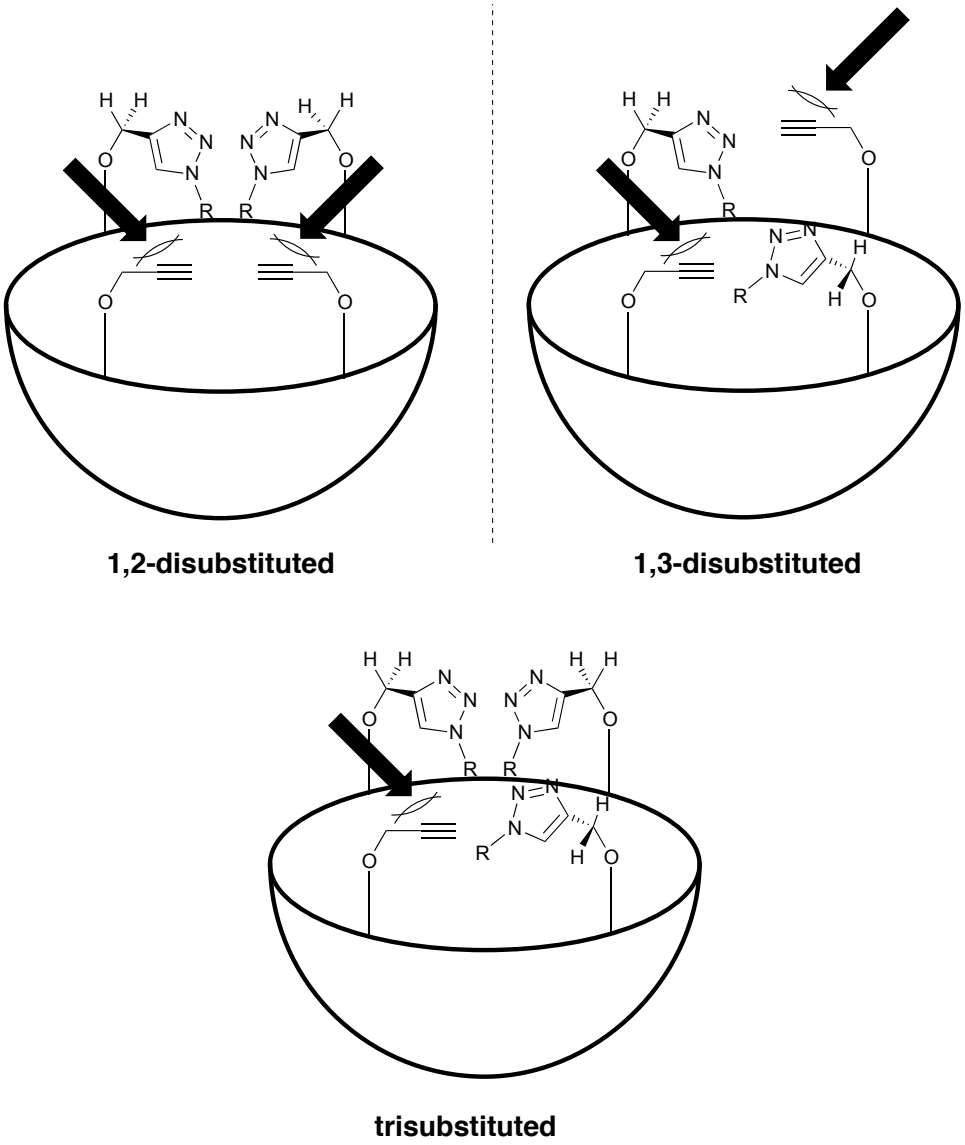
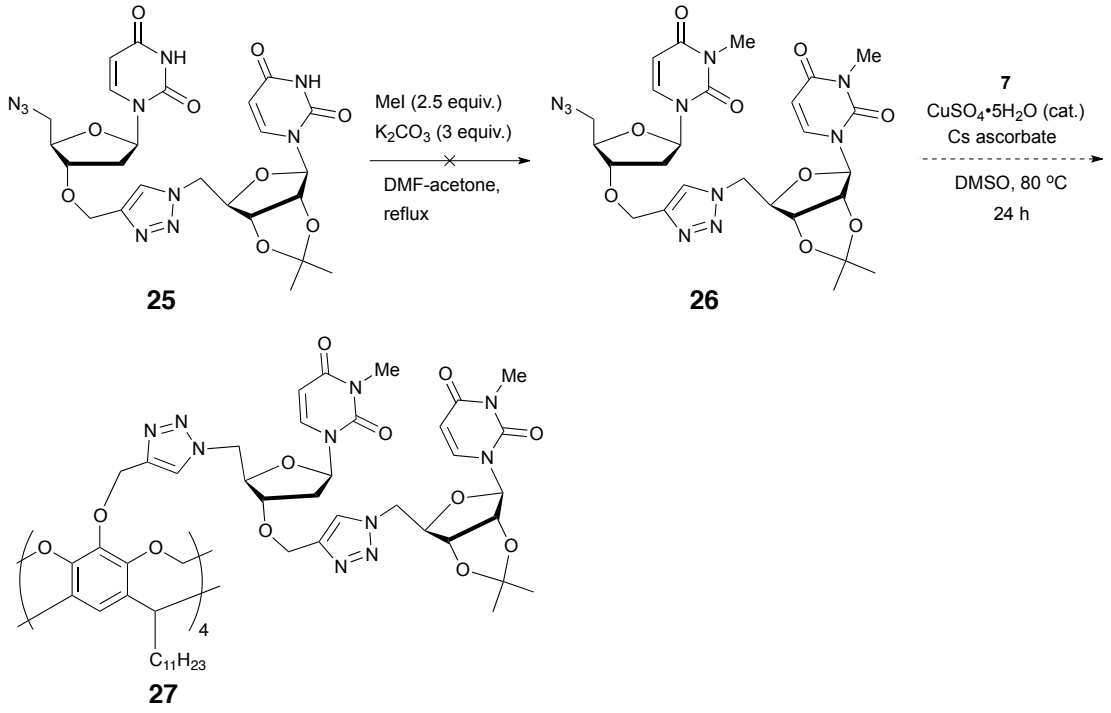


Figure 3.2 Schematic showing di- and trisubstituted conjugates sterically crowded at the cavitand rim, blocking approach of subsequent residues of 25. Monosubstituted conjugate not shown.

3.3.2 Synthesis of *N*-methylated conjugate

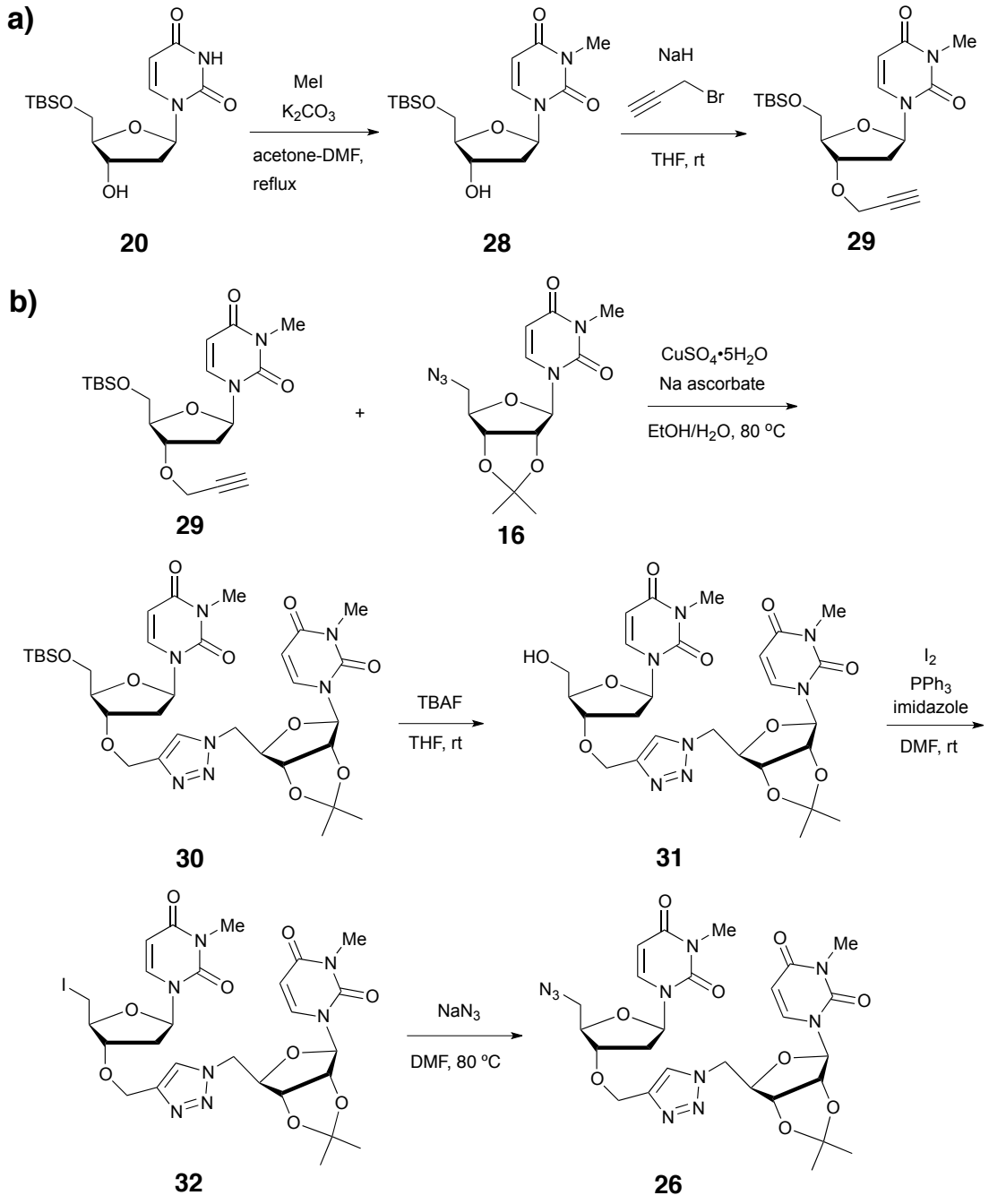
For subsequent control studies, the preparation of a conjugate with fully methylated imino positions to exclude hydrogen bonding was undertaken. Synthesis was expected to begin smoothly from dinucleoside **25** with a single methylation step via treatment with excess amounts

of iodomethane and K_2CO_3 base to provide the *N*-methylated derivative **26**, which would then be coupled with cavitand **7** to access conjugate **27** (Scheme 3.3). However, methylation of both imino positions on **25** in a single operation proved impossible. Reaction mixtures were found by LC-MS to contain only singly methylated product in addition to the starting material after 20 hours under refluxing conditions.



Scheme 3.3 Original synthetic plan for *N*-methylated conjugate 27.

Synthesis of **26** was therefore performed beginning with the *N*-methylation of **20** to afford **28**. Propargylation of the 3'-OH position provided propargyl ether **29** (Scheme 3.4a), which was then coupled with **16** in a Cu-catalyzed "click" reaction to give dinucleoside **30**. Subsequent transformations towards azide **26** were effected with identical methods to that of **25** (Scheme 3.4b).



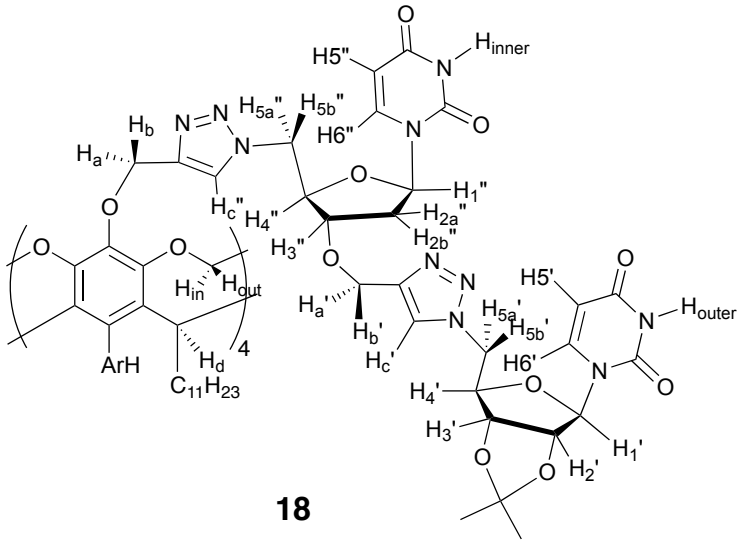
Scheme 3.4 a) Synthesis of propargyl ether 29. b) Synthesis of azide 26.

The Cu-catalyzed "click" reaction between cavitand **7** and azide **26** under the same reaction conditions for the synthesis of **18** proceeded cleanly to afford conjugate **27** in moderate yield (second step of Scheme 3.3), and its identity was confirmed by MALDI-TOF analysis.

3.4 NMR characterization studies

3.4.1 ^{1}H signal assignment

Significant ^{1}H signal broadening of **18** in $CDCl_3$ is observed at both room temperature and 55 ºC, and only ArH is resolved (Figure 3.5). This is indicative of intermolecular aggregation. Signal assignment was therefore performed in DMSO-d_6 with ^{1}H-^{1}H COSY. **27** was assigned in $CDCl_3$ and DMSO-d_6. ^{1}H and ^{1}H-^{1}H COSY spectra for both compounds are included in the Experimental Section. Figure 3.4 shows the condensed structure of **18** with all protons labeled. **27** differs only by both *N*-methyl groups and therefore is not shown.



18

Figure 3.3 Condensed structure of 18.

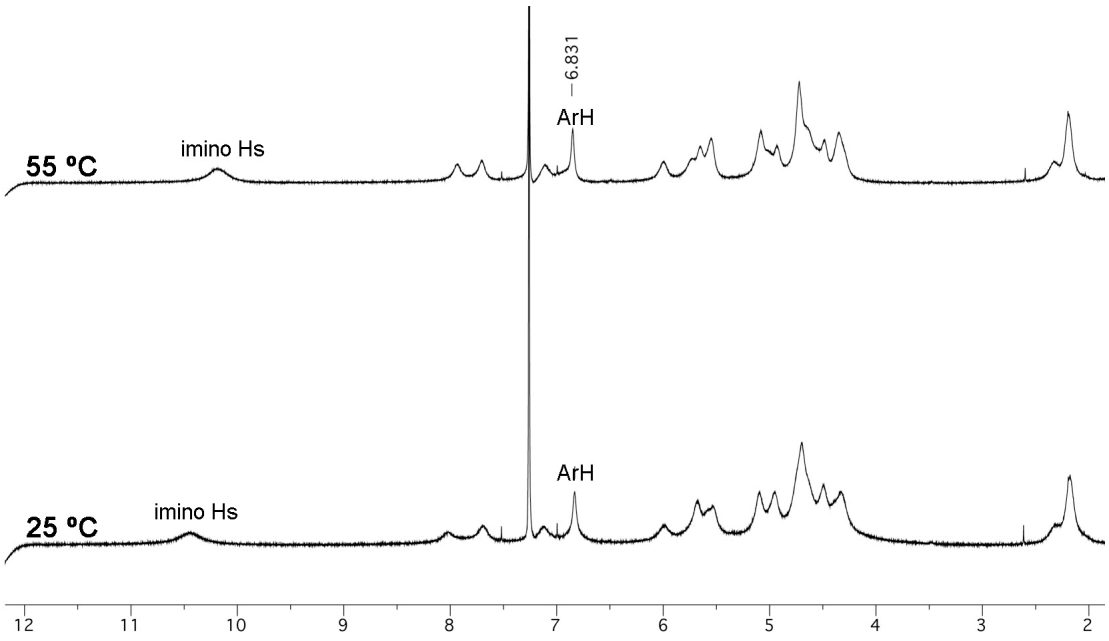


Figure 3.4 ^{1}H NMR spectra of 18 in $CDCl_3$ at 25 ºC and 55 ºC.

Signal overlap is widespread in the NMR spectra of **18** in DMSO-d_6 due to the similar chemical shifts of protons in both ribose rings as well as the linker methylene protons, necessitating group assignment. Nevertheless, important signals for structure determination remain resolved, such as the imino, H5 and H6 and H_c protons.

Imino protons (H_{inner} and H_{outer}) were identified by exchange with CD_3OD with observed quenching of both signals. In addition, weak COSY crosspeaks arising from W-couplings with the H5 protons (H5' and H5") provide further evidence of their identity (Figure 3.6). H_{outer} is assigned further downfield due to anisotropic deshielding. The H5' and H5" protons, which are overlapped, were in turn identified by COSY interactions with their respective H6 partners (H6' and H6") and the characteristic ^{3}J coupling constants (~ 8 Hz), calculated from the H6 doublets.

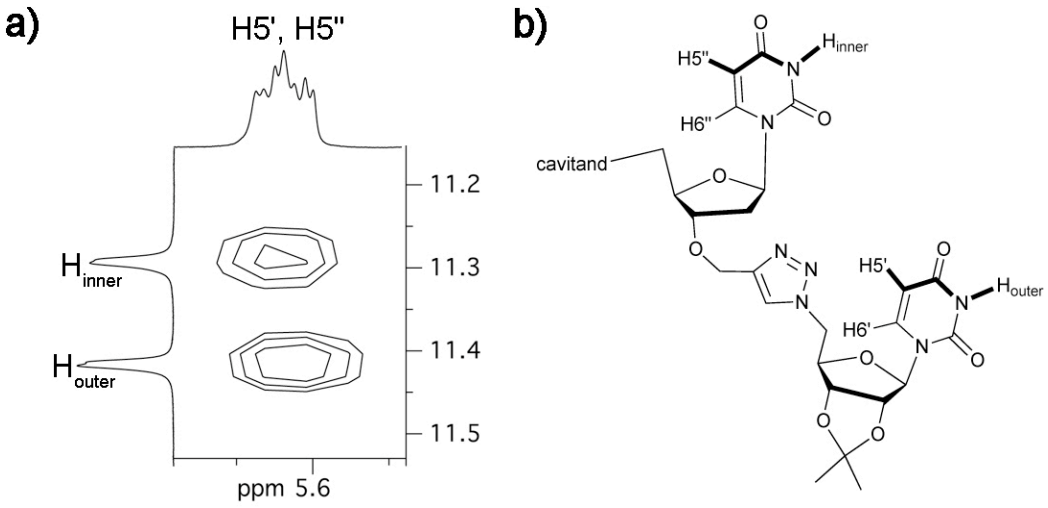


Figure 3.5 a) Portion of 400 MHz COSY spectrum of 18 in DMSO-d_6 showing W-coupling-based COSY crosspeaks between imino and H5 protons. b) W-coupling within uracil shown in bolded bonds.

All protons attached to the cavitand architecture were assigned by the appropriate COSY and NOESY correlations as applied to **1**.

3.4.2 Regiochemical analysis and signal assignment by NOESY

NOE spectra of **18** and **27** in DMSO-d_6 revealed strong correlations between ArH and H_d, as well as to the methylene -CH_2- 'feet' (Figure 3.7). It is noted that the actual H_d signal is overlapped with that of H_a and H_b, but since the spatial distance between ArH and H_a/H_b far exceeds that required for NOE interaction, the observed crosspeak corresponds to the ArH/H_d connection.

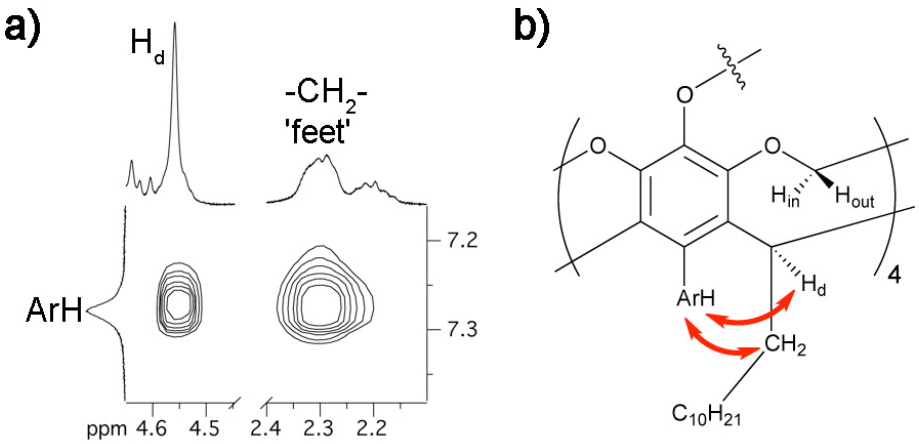


Figure 3.6 a) Portion of 400 MHz NOE spectrum of 18 in DMSO-d_6 showing crosspeaks between ArH and H_d and the –CH_2- 'feet' (27 exhibits similar correlations and is therefore not shown). b) NOE interactions of said protons.

Both triazole protons (H_c' and H_c") were identified by their NOE correlations with the respective methylene linkers (H_c': H_a'/H_b' and H_{5a}'/H_{5b}', H_c": H_a/H_b and H_{5a}"/H_{5b}") (Figure 3.8). These also confirm the 1,4-regiochemistry expected of the Cu-catalyzed "click" reaction in both triazole rings of **18** and **27**.

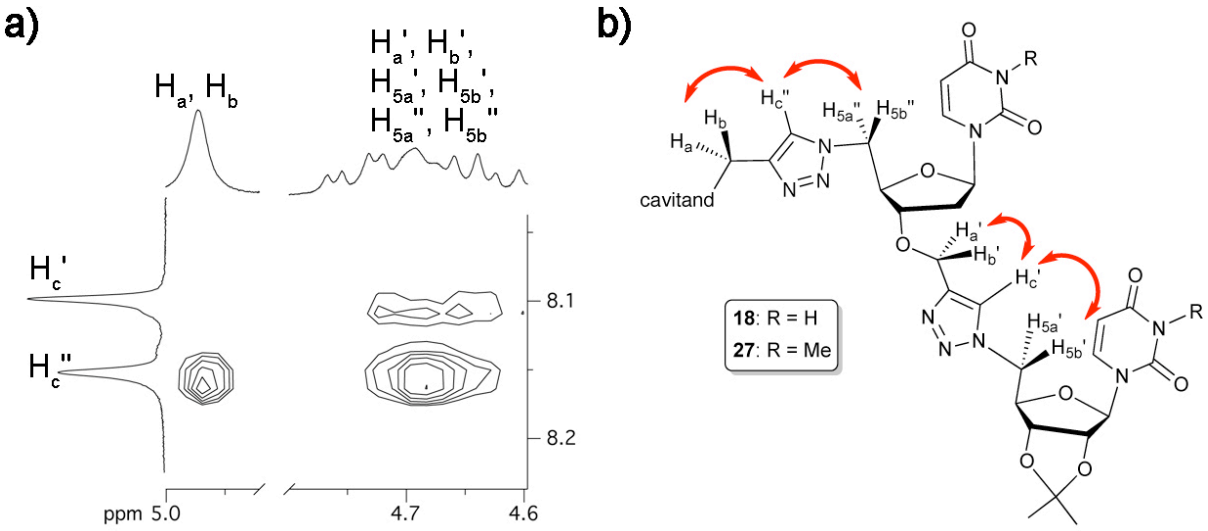


Figure 3.7 a) Portion of 400 Mhz NOE spectrum of 18 in DMSO-d_6 showing crosspeaks between triazole protons and flanking methylene groups (27 exhibits similar correlations and is therefore not shown). b) Expected NOE interactions of 1,4-regioisomers of 18.

3.4.3 Summary of ^{1}H signal assignments

Signal assignments for **18** and **27** are listed in Tables 3.1 through 3.3 with appropriate COSY and/or NOESY correlations used in their identifications included. Overlapped signals are group assigned and specified on the list.

Proton	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
H_{outer}	11.46	H5'	
H_{inner}	11.33	H5"	
H_c'	8.10		H_a'/H_b', H_{5a}'/H_{5b}'
H_c"	8.15		H_a/H_b, H_{5a}"/H_{5b}"
H6'	7.65	H5'	
H6"	7.57	H5"	
H5'	5.62	H6'	
H5"	5.62	H6"	
ArH	7.27		H_d
H_1'	5.75	H_2'	
H_1"	6.09	H_{2a}"/H_{2b}"	
H_2'	5.11	H_1', H_3'*	
H_{2a}"/H_{2b}"	2.25 – 2.29	H_{2b}"/H_{2a}"*	
H_3'	4.88	H_2', H_4'	
H_{out}	5.91	H_{in}	
H_a/H_b	4.96	H_b/H_a	
H_{5a}'/H_{5b}', H_{5a}"/H_{5b}", H_a'/H_b', H_d	4.55 – 4.76		
H_3", H_4', H_4", H_{in}	4.26 – 4.39		
-CH_2- 'feet'	2.15 – 2.22	H_d	
long chain 'feet', isopropylidene CH_3	1.20 – 1.41		
terminal CH_3	0.85		

*Correlation indistinguishable due to signal overlapping in the region.

Table 3.1 ^{1}H signal assignments of 18 in DMSO-d_6 at 25 ºC.

Proton	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.09		
N-Me	3.13		
H_c'	8.10		H_a'/H_b', H_{5a}'/H_{5b}'
H_c"	8.13		H_a/H_b, H_{5a}"/H_{5b}"
H6'	7.70	H5'	
H6"	7.61	H5"	
H5', H5", H_1'	5.73 – 5.79	H6', H6", H_2'	
ArH	7.26		H_d
H_1"	6.11	H_{2a}"/H_{2b}"	
H_2'	5.11	H_1', H_3'*	
H_3', H_a/H_b	4.89 – 4.96	H_2'/H_4'*, H_b/H_a	
H_{out}	5.88	H_{in}	
H_{5a}'/H_{5b}', H_{5a}"/H_{5b}", H_a'/H_b', H_d	4.55 – 4.77		
H_3", H_4', H_4", H_{in}	4.25 – 4.39		
H_{2a}"/H_{2b}", -CH_2- 'feet'	2.15 – 2.22	H_{2b}"/H_{2a}"*, H_d	
long chain 'feet', isopropylidene CH_3	1.22– 1.42		
terminal CH_3	0.83		

*Correlation indistinguishable due to signal overlapping in the region.

Table 3.2 ^{1}H signal assignments of 27 in DMSO-d_6 at 25 °C.

Proton	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.25		
N-Me	3.30		
H_c'	7.63		
H_c"	7.76		
H6'	7.12	H5'	H5'
H6"	7.02	H5"	H5"
H5', H5", H_{out}	5.68 – 5.76	H6', H6"	
ArH	6.82		
H_1'	5.53		
H_1"	6.13	H_{2a}"/H_{2b}"	
H_2', H_3', H_4', H_{5a}', H_{5b}', H_3", H_4", H_{5a}", H_{5b}", H_a, H_b, H_a', H_b', H_d, H_{in}	4.28 – 5.09		
H_{2a}"/H_{2b}", -CH_2- 'feet'	2.10 – 2.38	H_{2b}"/H_{2a}"	
long chain 'feet', isopropylidene CH_3	1.27– 1.40		
terminal CH_3	0.88		

Table 3.3 ^{1}H signal assignments of 27 in $CDCl_3$ at 25 ºC.

3.4.4 NMR solution structure elucidation

NOESY performed on **18** in DMSO-d_6 at 25 ºC revealed correlations between both the imino protons (H_{inner} and H_{outer}) and the H5 protons, suggesting the assembly of both the inner and outer U-quartets to form a quadruplex (Figure 3.9a). These crosspeaks are of noticeably different intensities, and the more intense crosspeak is assigned to H_{inner} (11.33 ppm) due to the expected tighter association of the inner U-quartet, capped by the more loosely bound outer quartet to yield a weaker crosspeak (H_{outer} = 11.46 ppm). In contrast, no crosspeaks were observed between the *N*-Me groups and H5 protons of **27** in DMSO-d_6, indicating the absence of quartets as expected due to the lack of imino protons.

Free dinucleoside **25** also shows no NOE correlation between its imino and H5 protons in DMSO-d_6, affirming the indispensability of the cavitand template in directing structure assembly. Water-mediated exchange between the imino protons is evident in the form of intense mutual crosspeaks as well as with residual water in the solvent (Figure 3.9b). Imino protons are labeled in Figure 3.9c. Neither of these correlations are found in the NOE spectrum of **18**, indicating that the imino protons therein are unexposed to the solvent, possible only in a closed quadruplex system, and that hydrogen bonding involving the imino protons is in effect, slowing their exchange.

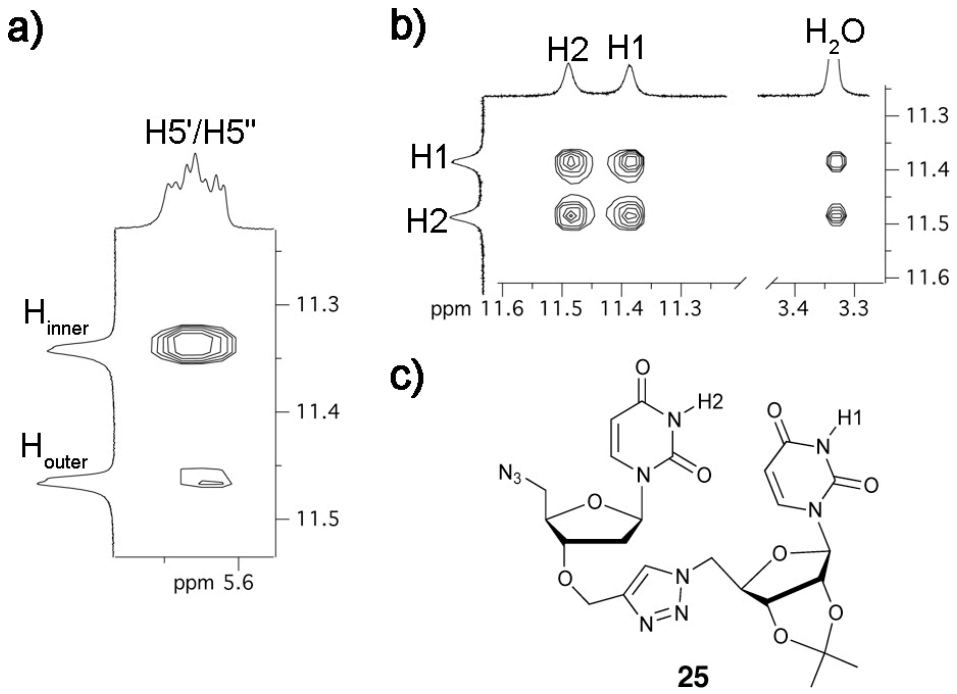


Figure 3.8 a) Portion of 400 MHz NOE spectrum of 18 in DMSO-d_6 showing crosspeaks between both imino protons and H5 protons. b) Portion of 400 Mhz NOE spectrum of 25 in DMSO-d_6 showing mutual imino crosspeaks and with residual H_2O in NMR solvent. c) Imino protons labeled on the structure of 25.

The assembly of hydrogen-bonded structures in DMSO is unusual, but not unprecedented. Base pairing and self-association of nucleosides in DMSO has been reported.[126-129] Stabilization of the system in **18** could arise not only from the energy of formation of eight hydrogen bonds (via the assembly of both quartets), but also from stacking of the uracil and triazole rings. This is evidenced by observed NOE correlations between H_c" and H6" as well as H_c' and H6" indicating close proximity of the inner triazole, inner uracil and outer triazole moieties (Figure 3.10a). The lack of a H_c'/H6' crosspeak suggests the non-planarity of the outer quartet, in turn giving rise to a more weakly associated quartet as described earlier in this section. Thus, a planar inner quartet is proposed which provides optimum hydrogen bond lengths according to the theoretical model, and the outer quartet 'dipped' inwards. This arrangement has

been observed in the solid state,[93] and has been assigned longer and thereby weaker hydrogen bonds by the theoretical model.[89] A visualization of the structure is provided in Figure 3.10b.

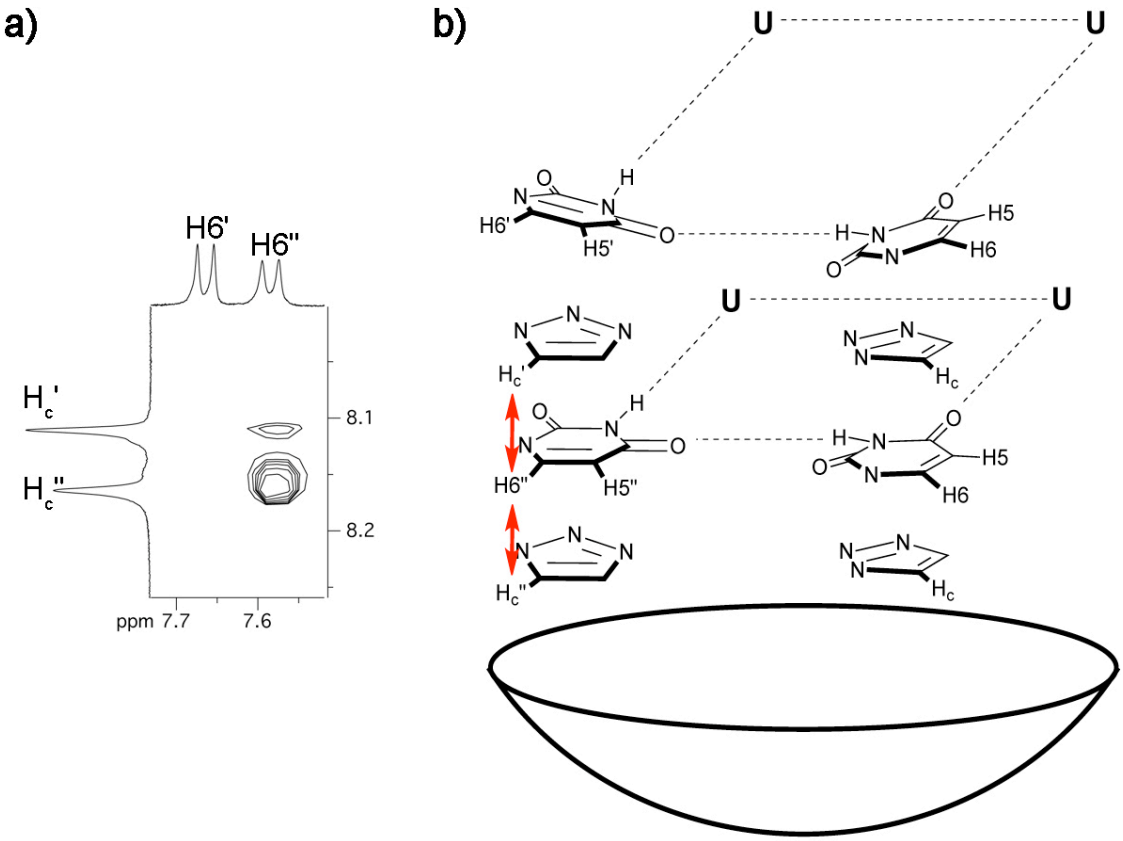


Figure 3.9 a) Portion of 400 MHz NOE spectrum of 18 in DMSO-d_6 showing crosspeaks between triazole protons and H6" suggestive of π-stacking. b) Illustration of proposed U-quadruplex of 18 showing planar and non-planar inner and outer U-quartets respectively. Triazole/H6 NOE connections are shown in red arrows.

S*yn* glycosidic bonds were confirmed at both ribose rings with the observation of NOE crosspeaks between H6' and H_1' as well as H6" and H_1" (Figure 3.11a). This is shown in Figure 3.11b.

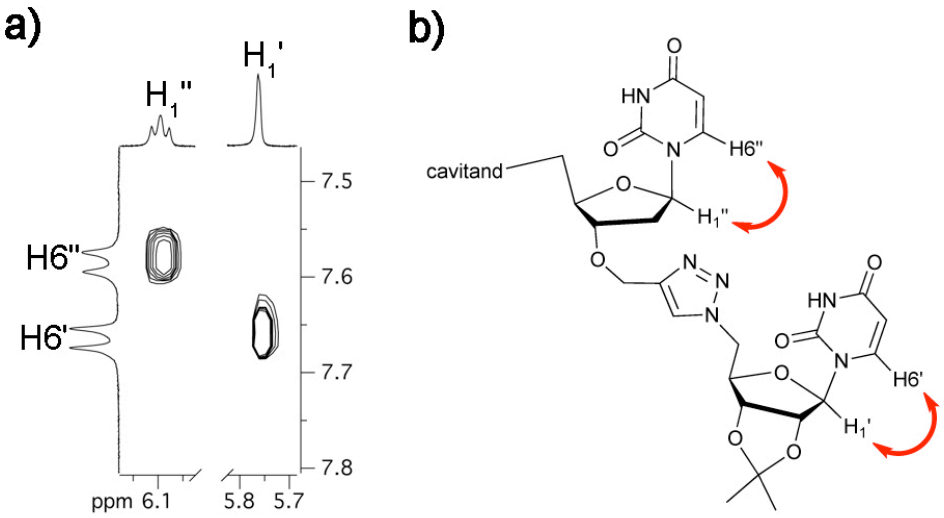


Figure 3.10 a) Portion of 400 MHz NOE spectrum of 18 showing H6'/H_1' and H6''/H_1'' crosspeaks indicative of *syn* glycosidic bonds. b) NOE correlations of the *syn* conformers shown in 18.

3.4.5 Variable-temperature NMR studies

The temperature dependence of the ^{1}H NMR spectrum of **18** in DMSO-d_6 was investigated (Figure 3.12). Pronounced upfield chemical shifts of both imino signals are observed at elevated temperatures with simultaneous signal broadening, suggesting the gradual disassembly of the U-quadruplex via hydrogen bond denaturation and an increasingly rapid exchange of the imino protons. Small upfield shifts of the H_c, H5 and H6 protons are also observed, indicating slight conformational changes of the triazole and uracil moieties with increasing temperature. The spectra show that the quadruplex retains some degree of its integrity over the observed temperature range.

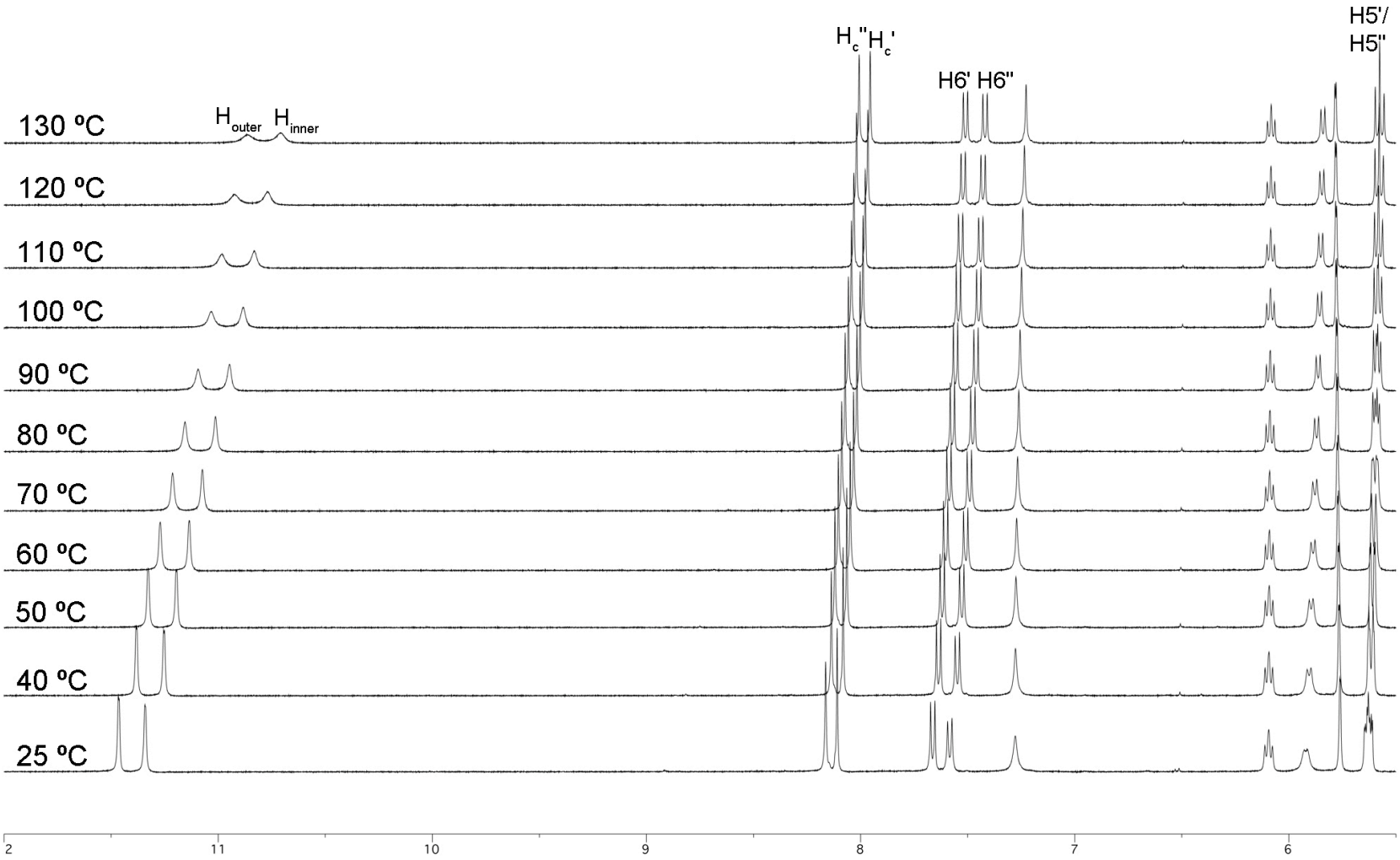


Figure 3.11 Variable-temperature ^{1}H NMR spectra of 18 in DMSO-d_6.

3.4.6 Diffusion NMR studies

DOSY NMR was performed on DMSO-d_6 and $CDCl_3$ solutions of **18** and **27** at identical concentration (2.7 mM) and temperature (25 ºC). The ArH proton was observed as it is the only isolated signal resolved for **18** in $CDCl_3$. *D* values were computed and generated by the SimFit algorithm. Measurements were taken thrice and the average values with standard error are reported in Table 3.4.

Solvent	**18**	**27**	Ratio
$CDCl_3$	1.95 ± 0.10	3.07 ± 0.10	0.64
DMSO-d_6	0.71 ± 0.10	0.73 ± 0.10	0.97

Table 3.4 ArH diffusion constants (*D*) of 18 **and** 27 **in $CDCl_3$ and DMSO-d_6 at 25 ºC (2.7 mM).**

Measured *D* values for both **18** and **27** in DMSO-d_6 are in close agreement, indicating an identical oligomeric state of the two compounds. Since no intermolecular aggregation is possible for **27**, both compounds are unimolecular in DMSO. In $CDCl_3$, the measured *D* value of **18** is 64% that of **27**, consistent with the theoretical ratio calculated for a termolecular with respect to a unimolecular system. This confirms the earlier hypothesis of the aggregation of **18** in $CDCl_3$.

3.5 Experimental section

3.5.1 General information

^{1}H NMR spectra were measured on a Bruker Avance 400 MHz in DMSO-d_6 [using DMSO (for 1H, δ = 2.50) as internal standard] or in $CDCl_3$ [using $CHCl_3$ (for ^{1}H, δ = 7.26) as internal standard]. Chemical shifts are reported in ppm from tetramethylsilane. ^{13}C NMR spectra were measured on a Bruker Avance 400 MHz spectrometer in DMSO-d_6 [using DMSO (for ^{13}C, δ = 39.5) as internal standard] or in $CDCl_3$ [using $CHCl_3$ (for ^{13}C, δ = 77.0) as internal standard]. The following abbreviations are used to explain the multiplicities: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, td = triplet of doublets, q = quartet, m = multiplet, br = broad. 2D NOESY spectra were acquired with t_{mix} = 800 ms and d1 = 1500 ms. COSY45 spectra were acquired with d1 = 1500 ms. DOSY experiments were carried out on a Bruker Avance 400inv spectrometer equipped with a 5mm BBI Z-gradient probe (inverse broadband probe with z-gradient coil). All measurements were performed using a BPLED gradient pulse sequence (ledbpgp2s). The length of the diffusion gradient was optimized for each sample to obtain at least 95% signal attenuation due to diffusion. Δ and δ values respectively were found to be 65 ms and 5 ms (for **18** and **27** in $CDCl_3$), 65 ms and 10 ms (for **18** in DMSO-d_6), 65 ms and 10.4 ms (for **27** in DMSO-d_6). Eddy current (t_e) was set at 5 ms. All measurements were taken at 298K with sample concentrations of 2.7 mM. Diffusion coefficients were generated using the SimFit function on the XWinNMR software. MALDI-TOF mass analyses were performed on a Bruker Biflex IV spectrometer in the reflectron mode using 2,5-dihydroxybenzoic acid (DHB) as the matrix. Flash column chromatography was performed using Silicycle 60 silica gel and eluting

solvents were used directly from their commercial bottles. Solvents and reagents for reactions were purchased commercially and used without further purification.

3.5.2 Synthesis of conjugate 18

3.5.2.1 Preparation of 5'-OH dinucleoside 23

To a stirred solution of **22** (1.31 g, 1.90 mmol) in THF (19 mL) was added TBAF (2.3 mL of a 1.0 M solution in THF, 2.28 mmol), and the mixture was stirred for 1 h. The solvent was then removed *in vacuo* and the residue re-dissolved in $CHCl_3$. The solvent was removed *in vacuo* once again, and the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 9 : 1 – $CHCl_3$: EtOH 4 : 1) to afford **23** in 96% yield (1.04 g, 1.81 mmol).

Compound 23: White solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.27 (3H, s, isopropylidene CH_3), 1.46 (3H, s, isopropylidene CH_3), 2.07 – 2.14 (1H, m, H_{2b}"), 2.25 – 2.32 (1H, m, H_{2a}"), 3.56 – 3.58 (2H, m, H_{5a}" and H_{5b}"), 3.97 – 4.00 (1H, m, H_4"), 4.19 – 4.20 (1H, m, H_3"), 4.35 – 4.38 (1H, m, H_4'), 4.56 (2H, s, H_a' and H_b'), 4.65 (1H, dd, J = 7.6, 14.0 Hz, H_{5b}'), 4.75 (1H, dd, J = 5.2, 14.0 Hz, H_{5a}'), 4.88 (1H, dd, J = 4.0, 6.4 Hz, H_3'), 5.07 (1H, t, J = 5.2 Hz, hydroxyl OH), 5.12 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.63 (2H, m, H5' and H5"), 5.76 (1H, d, J = 1.6 Hz, H_1'), 6.09 (1H, dd, J = 5.6, 8.0 Hz, H_1"), 7.66 (1H, d, J = 8.4 Hz, H6'), 7.84 (1H, d, J = 8.0 Hz, H6"), 8.13 (1H, s, H_c'), 11.30 (1H, s, H1), 11.46 (1H, s, H2); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.3, 162.1, 149.5, 149.3, 142.9, 142.6, 139.4, 123.6, 112.4, 100.95, 100.89, 92.4, 84.4, 83.8, 83.2,

82.7, 80.5, 78.1, 60.8, 60.5, 50.3, 35.6, 25.9, 24.1; ESIHRMS: Found: *m/z* 598.1882. Calcd for $C_{24}H_{29}N_7O_{10}Na$: $(M+Na)^+$ 598.1874.

3.5.2.2 Preparation of 5'-I dinucleoside 24

To a stirred solution of **4** (1.044 g, 1.814 mmol) in DMF (8 mL) was added PPh_3 (1.571 g, 6 mmol), imidazole (0.82 g, 12.04 mmol) and I_2 (1.456 g, 5.735 mmol, 4 portions over 5 mins) in that order. The reaction was stirred at room temperature for 4 h, whereupon the solvent was removed *in vacuo*. H_2O (10 mL) was added to the residue and the mixture was extracted with $CHCl_3$ (3 x 50 mL). The combined organic extracts were washed with brine and dried over anhydrous $MgSO_4$. After evaporation of the solvent, the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 19 : 1 – $CHCl_3$: EtOH 9 : 1) to afford **5** in 81% yield (1.007g, 1.469 mmol).

Compound 24: White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.30 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 2.26 – 2.39 (2H, m, H_{2a}" and H_{2b}"), 3.42 (1H, dd, *J* = 6.4, 10.4 Hz, H_{5a}" or H_{5b}"?), 3.51 (1H, dd, *J* = 6.8, 10.4 Hz, H_{5a}" or H_{5b}"?), 4.07 (1H, td, *J* = 2.4, 6.8, 6.8 Hz, H_4"), 4.15 – 4.18 (1H, m, H_3"), 4.37 – 4.41 (1H, m, H_4'), 4.62 – 4.67 (3H, m, H_a', H_b', H_{5b}'), 4.77 (1H, dd, *J* = 4.8, 14.0 Hz, H_{5a}'), 4.90 (1H, dd, *J* = 4.0, 6.4 Hz, H_3'), 5.15 (1H, dd, *J* = 1.6, 6.4 Hz, H_2'), 5.65 (1H, dd, *J* = 2.0, 7.6 Hz, H5'), 5.70 (1H, dd, *J* = 2.0, 8.0 Hz, H5"), 5.79 (1H, d, *J* = 2.0 Hz, H_1'), 6.15 (1H, dd, *J* = 6.0, 8.4 Hz, H_1"), 7.69 – 7.73 (2H, m, H6' and H6"), 8.18 (1H, s, H_c'), 11.38 (1H, d, *J* = 2.0 Hz, H1), 11.48 (1H, d, *J* = 2.0 Hz, H2); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.3, 162.9, 150.4, 150.3, 143.63, 143.59, 140.6, 124.7, 113.4,

102.2, 101.9, 93.4, 85.4, 84.7, 83.7, 83.1, 81.5, 80.8, 61.9, 51.3, 35.2, 26.9, 25.1, 7.6; ESIHRMS: Found: *m/z* 708.0899. Calcd for $C_{24}H_{28}N_7O_9NaI$: $(M+Na)^+$ 708.0891.

3.5.2.3 Preparation of 5'-N_3 dinucleoside 25

To a stirred solution of **5** (482.9 mg, 0.7045 mmol) in DMF (2.5 mL) was added NaN_3 (91.6 mg, 1.409 mmol). The reaction was then stirred at 80 ºC for 4 h, whereupon the solvent was removed *in vacuo*. H_2O (10 mL) was added to the residue and the mixture was extracted with $CHCl_3$ (3 x 50 mL). The combined organic extracts were washed with brine and dried over anhydrous $MgSO_4$. After evaporation of the solvent, the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 19 : 1 – $CHCl_3$: EtOH 9 : 1) to provide **2** in 83% yield (352 mg, 0.586 mmol).

Compound 25: White foamy solid; ^1H (400 MHz, DMSO-d_6) δ(ppm) 1.28 (3H, s, isopropylidene CH_3), 1.46 (3H, s, isopropylidene CH_3), 2.27 – 2.32 (2H, m, H_{2a}" and H_{2b}"), 3.52 – 3.61 (2H, m, H_{5a}" and H_{5b}"), 4.04 – 4.07 (1H, m, H_4"), 4.14 – 4.17 (1H, m, H_3"), 4.35 – 4.39 (1H, m, H_4'), 4.58 (2H, s, H_a' and H_b'), 4.63 (1H, dd, J = 8.0, 14.0 Hz, H_{5b}'), 4.75 (1H, dd, J = 5.2, 14.0 Hz, H_{5a}'), 4.88 (1H, dd, J = 4.0, 6.4 Hz, H_3'), 5.12 (1H, dd, J = 1.2, 6.0 Hz, H_2'), 5.65 (2H, m, H5' and H5"), 5.77 (1H, d, J = 1.2 Hz, H_1'), 6.12 (1H, t, J = 6.8 Hz, H_1"), 7.67 – 7.70 (2H, m, H6' and H6"), 8.15 (1H, s, H_c'), 11.37 (1H, s, H1), 11.47 (1H, s, H2); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.3, 162.9, 150.4, 150.3, 143.7, 143.6, 140.7, 124.6, 113.4, 102.2, 101.9, 93.4, 85.4, 84.5, 83.7, 82.2, 81.5, 78.8, 62.0, 51.8, 51.3, 35.4, 26.8, 25.1; ESIHRMS: Found: *m/z* 623.1945. Calcd for $C_{24}H_{28}N_{10}O_9Na$: $(M+Na)^+$ 623.1938.

3.5.2.4 Preparation of conjugate 18 via the "click" reaction

To a stirred solution of cavitand **3** (30.7 mg, 0.0224 mmol) and nucleoside **2** (60.6 mg, 0.1008 mmol) in argon-purged DMSO (2.7 mL) was added a solution of copper(II) sulfate pentahydrate (75 μL, 0.04 M in argon-purged Milli-Q water, 0.00298 mmol) followed by a solution of cesium ascorbate, freshly prepared from the stoichiometric reaction between cesium carbonate and ascorbic acid (75 μL, 0.4 M in argon-purged Milli-Q water, 0.0298 mmol). The reaction was stirred at 80 °C for 24 h. The solvent was then removed *in vacuo* and the residue suspended in water. A few drops of ammonium hydroxide were added to remove the copper catalyst and the mixture was suction filtered. The residue, crude **1**, was washed with deionized water and allowed to suction dry, whereupon it was purified via step gradient flash chromatography ($CHCl_3$: MeOH 9 : 1 – $CHCl_3$: MeOH 85 : 15) to afford conjugate **1** in 36% yield (30.2 mg, 0.008 mmol).

Compound 18: White glassy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 0.84 (12H, t, J = 7.2 Hz, CH_3 feet), 1.20 – 1.41 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 2.15 – 2.22 (8H, m, CH_2 feet), 2.25 – 2.29 (8H, m, H_{2a}" and H_{2b}"), 4.26 – 4.39 (16H, m, H_{in}, H_3", H_4', H_4"), 4.55 – 4.76 (28H, m, H_d, H_a', H_b', H_{5a}', H_{5b}', H_{5a}", H_{5b}"), 4.88 (4H, dd, J = 4.0, 6.4 Hz, H_3'), 4.96 (8H, s, H_a and H_b), 5.11 (4H, dd, J = 0.8, 6.4 Hz, H_2'), 5.60 – 5.64 (8H, m, H5' and H5"), 5.75 (4H, s, H_1'), 5.91 (4H, d, J = 5.6 Hz, H_{out}), 6.09 (4H, t, J = 6.8 Hz, H_1"), 7.27 (4H, s, ArH), 7.57 (4H, d, J = 8.0 Hz, H6"), 7.65 (4H, d, J = 8.0 Hz, H6'), 8.10 (4H, s, H_c'), 8.15 (4H, s, H_c"), 11.33 (4H, s, H_{inner}), 11.46 (4H, s, H_{outer}); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.9, 163.5, 150.94, 150.91, 148.1, 144.4, 144.2, 144.1, 141.2, 139.4, 125.72, 125.69, 125.2, 114.0, 102.7, 102.5, 94.0, 86.0, 85.4, 84.3, 82.1, 79.6, 62.7, 51.94, 51.88, 35.9, 31.9, 30.0, 29.9, 29.7, 29.4, 28.4, 27.4,

25.7, 22.7, 14.5; MS(MALDI-TOF): Found: *m/z* 3776.1. Calcd for $C_{184}H_{233}N_{40}O_{48}$: $(M+H)^+$ 3773.1.

3.5.3 Synthesis of conjugate 27 (procedure similar to that of 18)

3.5.3.1 Preparation of 5'-OH dinucleoside 31

The procedure for the preparation of **31** was similar to that of **23**.

Compound 31: 96% yield; White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.28 (3H, s, isopropylidene CH_3), 1.47 (3H, s, isopropylidene CH_3), 2.08 – 2.15 (1H, m, H_{2b}"), 2.28 – 2.33 (1H, m, H_{2a}"), 3.14 (3H, s, N-Me), 3.15 (3H, s, N-Me), 3.52 – 3.61 (2H, m, H_{5a}" and H_{5b}"), 3.98 – 4.00 (1H, m, H_4"), 4.19 – 4.21 (1H, m, H_3"), 4.36 – 4.40 (1H, m, H_4'), 4.56 (2H, s, H_a' and H_b'), 4.63 (1H, dd, *J* = 7.6, 14.0 Hz, H_{5b}'), 4.75 (1H, dd, *J* = 4.8, 14.0 Hz, H_{5a}'), 4.91 (1H, dd, *J* = 4.0, 6.4 Hz, H_3'), 5.10 (1H, t, *J* = 5.2 Hz, hydroxyl OH), 5.14 (1H, dd, *J* = 1.6, 6.4 Hz, H_2'), 5.76 (1H, d, *J* = 2.4 Hz, H5'), 5.78 (1H, d, *J* = 2.4 Hz, H5"), 5.80 (1H, d, *J* = 1.6 Hz, H_1'), 6.12 (1H, dd, *J* = 6.0, 8.0 Hz, H_1"), 7.73 (1H, d, *J* = 8.0 Hz, H6'), 7.90 (1H, d, *J* = 8.4 Hz, H6"), 8.13 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.0, 161.8, 150.4, 150.3, 143.5, 141.7, 138.4, 124.5, 113.2, 100.7, 94.1, 85.4, 85.1, 84.7, 83.54, 83.50, 81.2, 78.7, 61.6, 61.1, 51.0, 36.7, 27.0, 26.9, 26.6, 24.8; ESIHRMS: Found: *m/z* 626.2171. Calcd for $C_{26}H_{33}N_7O_{10}Na$: $(M+Na)^+$ 626.2187.

3.5.3.2 Preparation of 5'-I dinucleoside 32

The procedure for the preparation of **32** was similar to that of **24**.

Compound 32: 82% yield; White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 2.28 – 2.39 (2H, m, H_{2a}" and H_{2b}"), 3.16

(3H, s, N-Me), 3.17 (3H, s, N-Me), 3.42 (1H, dd, J = 6.4, 10.4 Hz, H_{5a}" or H_{5b}"), 3.51 (1H, dd, J = 6.4, 10.4 Hz, H_{5a}" or H_{5b}"), 4.09 (1H, td, J = 2.0, 6.4, 6.4 Hz, H_4"), 4.15 – 4.18 (1H, m, H_3"), 4.38 – 4.43 (1H, m, H_4'), 4.62 – 4.68 (3H, m, H_a', H_b', H_{5b}'), 4.78 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.93 (1H, dd, J = 4.0, 6.4 Hz, H_3'), 5.16 (1H, dd, J = 2.0, 6.8 Hz, H_2'), 5.78 – 5.84 (3H, m, H5', H5", H_1'), 6.18 (1H, dd, J = 6.4, 7.6 Hz, H_1"), 7.74 (1H, d, J = 2.0 Hz, H6'), 7.76 (1H, d, J = 2.0 Hz, H6"), 8.17 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.2, 161.9, 150.64, 150.55, 143.6, 141.9, 138.9, 124.8, 113.4, 101.2, 100.9, 94.4, 85.8, 85.7, 83.7, 83.3, 81.4, 80.8, 61.9, 51.3, 35.5, 27.3, 27.2, 26.9, 25.1, 7.4; ESIHRMS: Found: *m/z* 714.1415. Calcd for $C_{26}H_{33}N_7O_9I$: $(M+H)^+$ 714.1385.

3.5.3.3 Preparation of 5'-N_3 dinucleoside 26

The procedure for the preparation of **26** was similar to that of **25**.

Compound 26: 91% yield; White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 2.30 – 2.33 (2H, m, H_{2a}" and H_{2b}"), 3.16 (3H, s, N-Me), 3.17 (3H, s, N-Me), 3.54 – 3.64 (2H, m, H_{5a}" and H_{5b}"), 4.07 – 4.10 (1H, m, H_4"), 4.16 – 4.19 (1H, m, H_3"), 4.38 – 4.42 (1H, m, H_4'), 4.59 (2H, s, H_a' and H_b'), 4.65 (1H, dd, J = 7.6, 14.0 Hz, H_{5b}'), 4.78 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.93 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.16 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.78 – 5.84 (3H, m, H_1', H5', H5"), 5.77 (1H, d, J = 1.2 Hz, H_1'), 6.17 (1H, t, J = 6.8 Hz, H_1"), 7.74 (1H, d, J = 2.0 Hz, H6'), 7.76 (1H, d, J = 2.0 Hz, H6"), 8.16 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.2, 161.9, 150.7, 150.6, 143.6, 141.9, 139.0, 124.8, 113.4, 101.1, 100.9, 94.4, 85.7, 83.7, 82.3, 81.4, 78.8, 62.1, 51.8, 51.3, 35.7, 27.3, 27.2, 26.8, 25.1; ESIHRMS: Found: *m/z* 629.2438. Calcd for $C_{26}H_{33}N_{10}O_9$: $(M+H)^+$ 629.2432.

3.5.3.4 Preparation of conjugate 27

The procedure for the preparation of **27** was similar to that of **18**.

Compound 27: 36% yield; White glassy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 0.83 (12H, t, J = 6.4 Hz, CH_3 feet), 1.22 – 1.42 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 2.22 – 2.29 (16H, m, CH_2 feet, H_{2a}", H_{2b}"), 3.09 (3H, s, N-Me), 3.13 (3H, s, N-Me), 4.25 – 4.39 (16H, m, H_{in}, H_3", H_4', H_4"), 4.55 – 4.77 (28H, m, H_d, H_a', H_b', H_{5a}', H_{5b}', H_{5a}", H_{5b}"), 4.89 – 4.96 (12H, m, H_3', H_a, H_b), 5.11 (4H, dd, J = 0.8, 6.4 Hz, H_2'), 5.73 – 5.79 (12H, m, H_1', H5', H5"), 5.88 (4H, d, J = 6.0 Hz, H_{out}), 6.11 (4H, t, J = 6.8 Hz, H_1"), 7.26 (4H, s, ArH), 7.61 (4H, d, J = 8.0 Hz, H36"), 7.70 (4H, d, J = 8.0 Hz, H6'), 8.10 (4H, s, H_c'), 8.13 (4H, s, H_c"); ^{1}H (400 MHz, $CDCl_3$) δ(ppm) 0.88 (12H, t, J = 6.4 Hz, CH_3 feet), 1.27 – 1.40 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 2.10 – 2.38 (12H, m, CH_2 feet, H_{2a}", H_{2b}"), 3.25 (12H, s, N-Me), 3.30 (12H, s, N-Me), 4.28 – 5.09 (60H, m, H_2', H_3', H_4', H_{5a}', H_{5b}', H_3", H_4", H_{5a}", H_{5b}", H_a, H_b, H_a', H_b', H_d, H_{in}), 5.53 (4H, s, H_1'), 5.68 – 5.76 (12H, m, H_{out}, H5', H5"), 6.13 (4H, t, J = 6.0 Hz, H_1"), 6.82 (4H, s, ArH), 7.02 (4H, d, J = 8.0 Hz, H6"), 7.12 (4H, d, J = 8.0 Hz, H6'), 7.63 (4H, s, H_c'), 7.76 (4H, s, H_c"); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.2, 161.9, 150.5, 147.4, 143.8, 143.6, 143.5, 141.9, 138.9, 138.8, 125.1, 124.7, 113.4, 101.1, 100.9, 94.4, 86.0, 85.6, 83.8, 81.9, 81.4, 78.9, 66.63, 66.60, 62.1, 51.3, 37.0, 35.5, 31.4, 29.33, 29.27, 29.11, 29.08, 28.8, 27.7, 27.2, 27.1, 26.8, 25.0, 22.1, 13.9; ^{13}C (100 MHz, $CDCl_3$) δ(ppm) 162.70, 162.65, 151.0, 148.2, 145.4, 144.3, 144.2, 141.3, 139.2, 139.1, 138.2, 124.8, 124.3, 114.8, 102.3, 99.7, 97.2, 87.3, 86.6, 84.6, 82.3, 82.2, 78.9, 77.4 (overlapped), 67.4, 63.3, 52.2, 51.5, 37.2, 36.9, 32.1, 30.2, 30.1, 30.0, 29.92, 29.90, 29.85, 29.6, 22.9, 14.3, 1.2; MS(MALDI-TOF): Found: *m/z* 3887.8. Calcd for $C_{192}H_{249}N_{40}O_{48}$: $(M+H)^+$ 3885.3.

3.5.4 Supplementary ^{1}H and ^{1}H-^{1}H COSY spectra

Spectra begin on the following page.

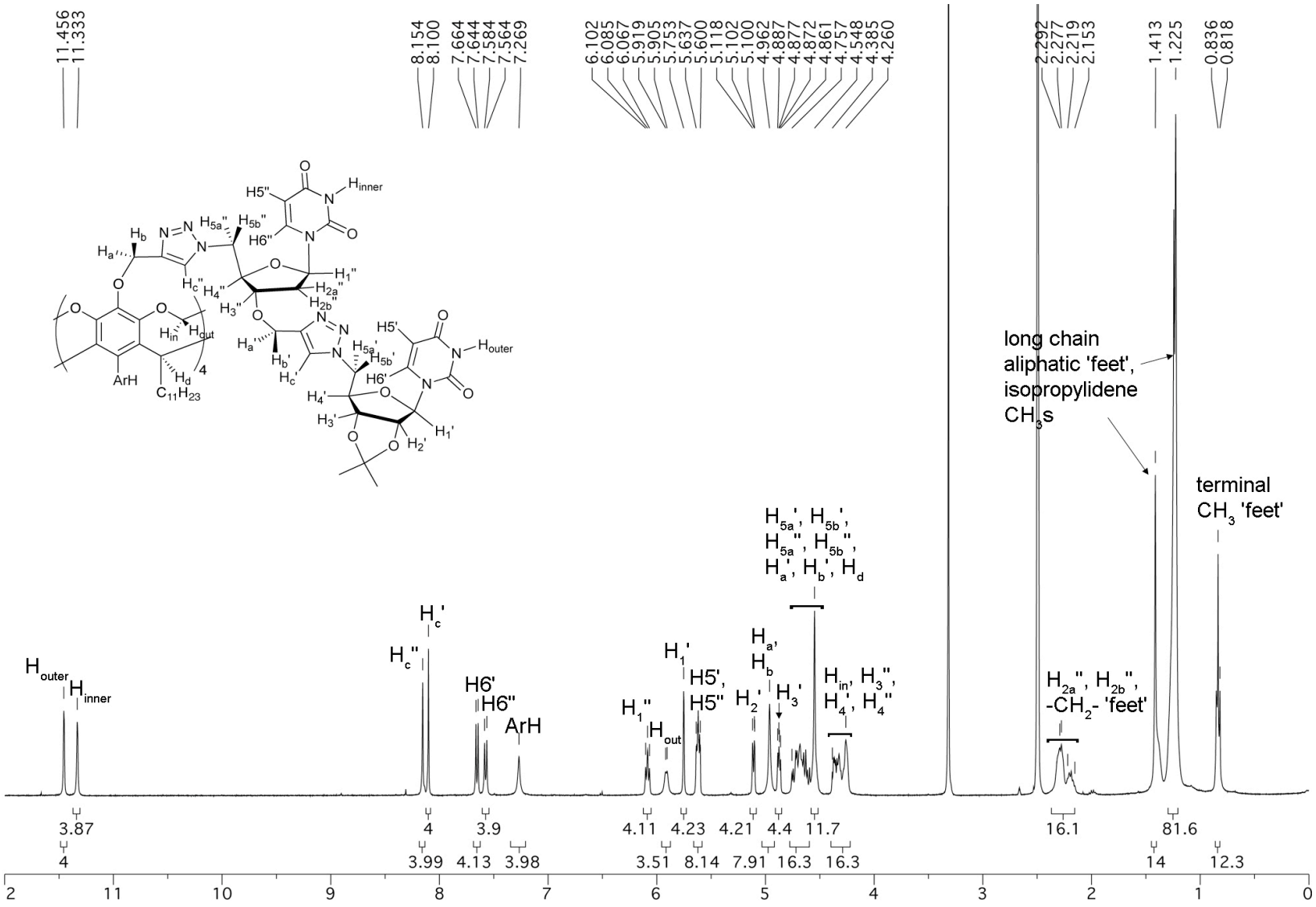


Figure 3.12 ^{1}H NMR spectrum of 18 in DMSO-d_6 at 25 ºC.

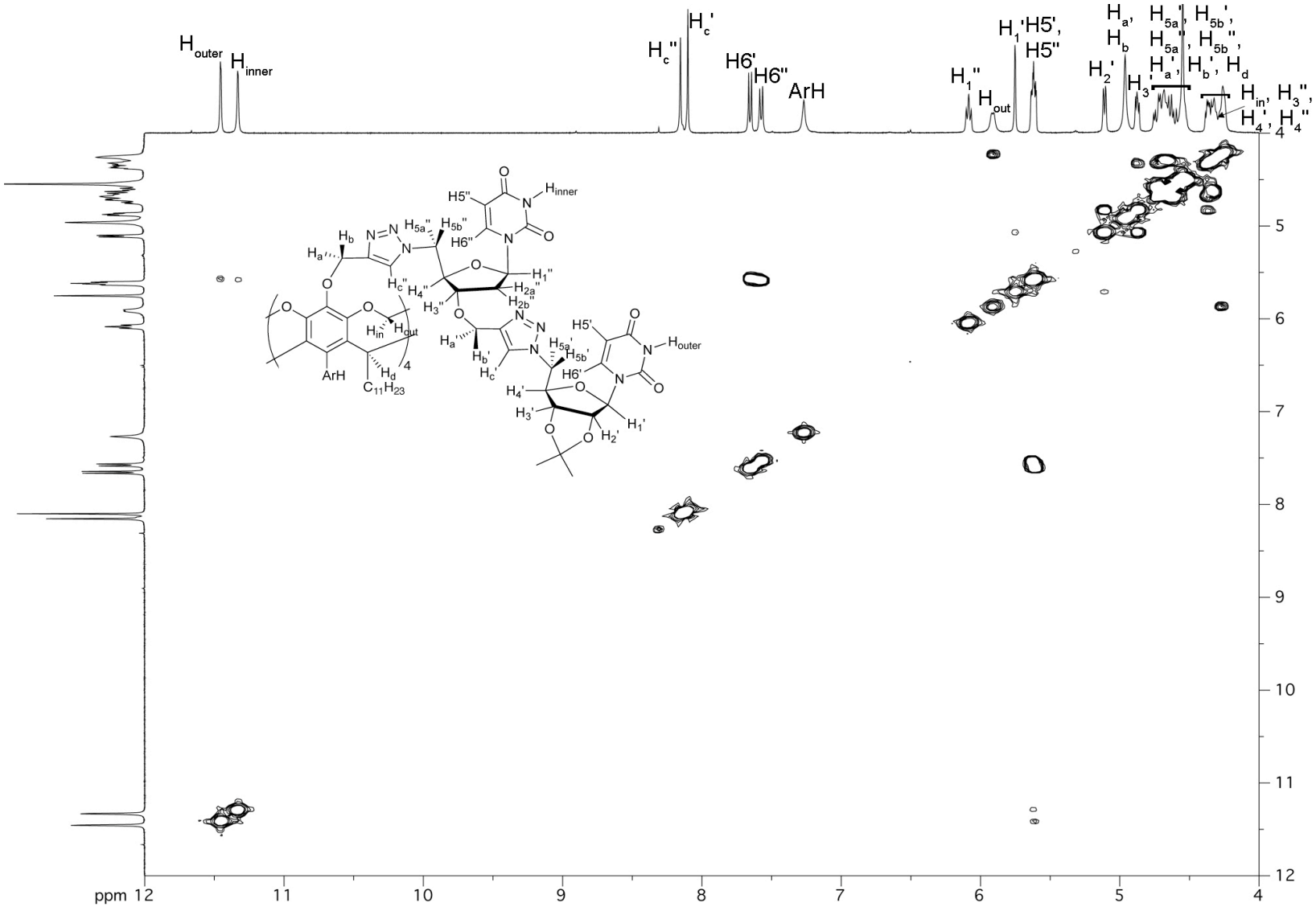


Figure 3.13 ^{1}H-^{1}H COSY spectrum of 18 in DMSO-d_6 at 25 ºC.

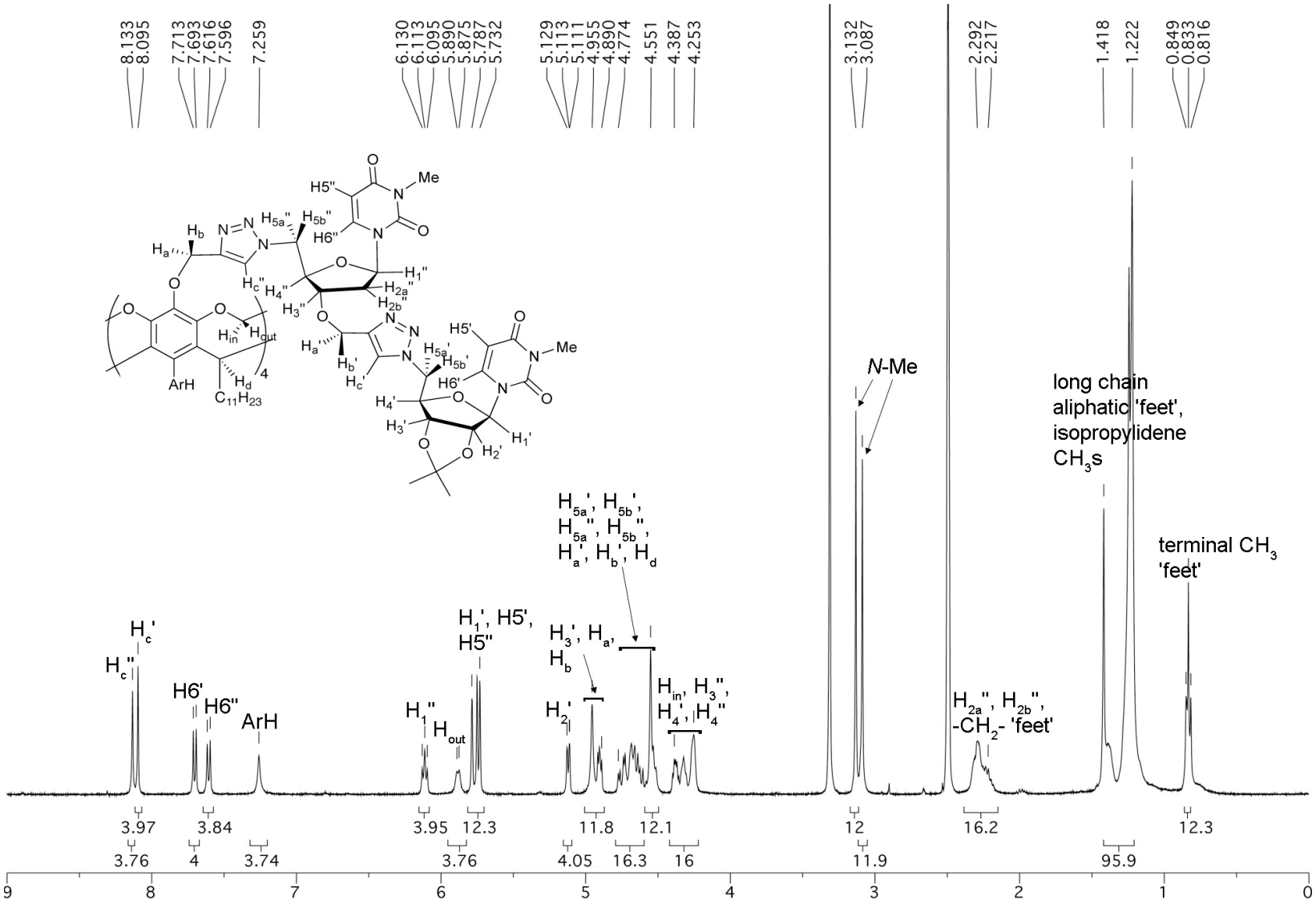


Figure 3.14 ^{1}H NMR spectrum of 27 in DMSO-d_6 at 25 ºC.

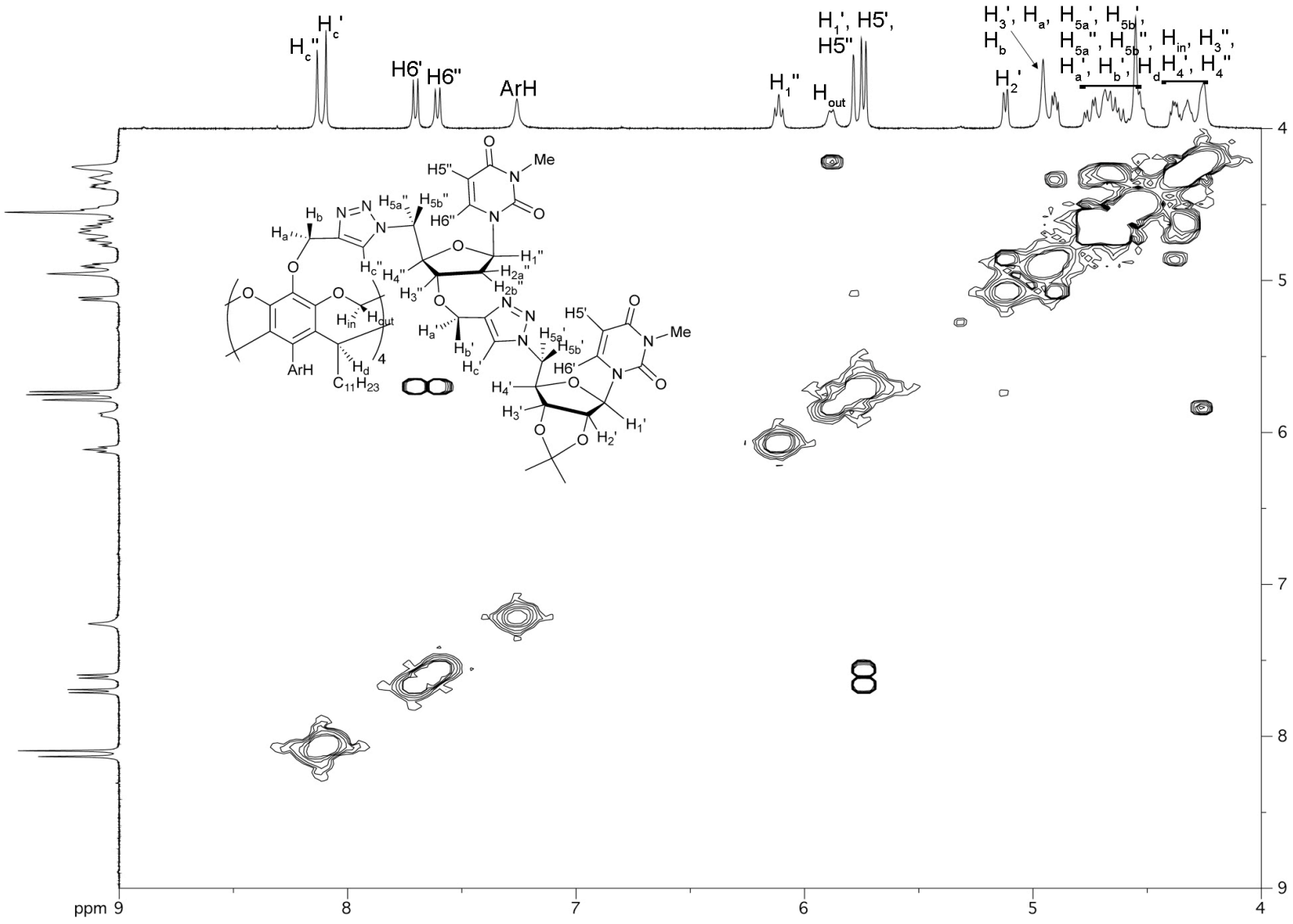


Figure 3.15 ^{1}H-^{1}H COSY spectrum of 27 in DMSO-d_6 at 25 ºC.

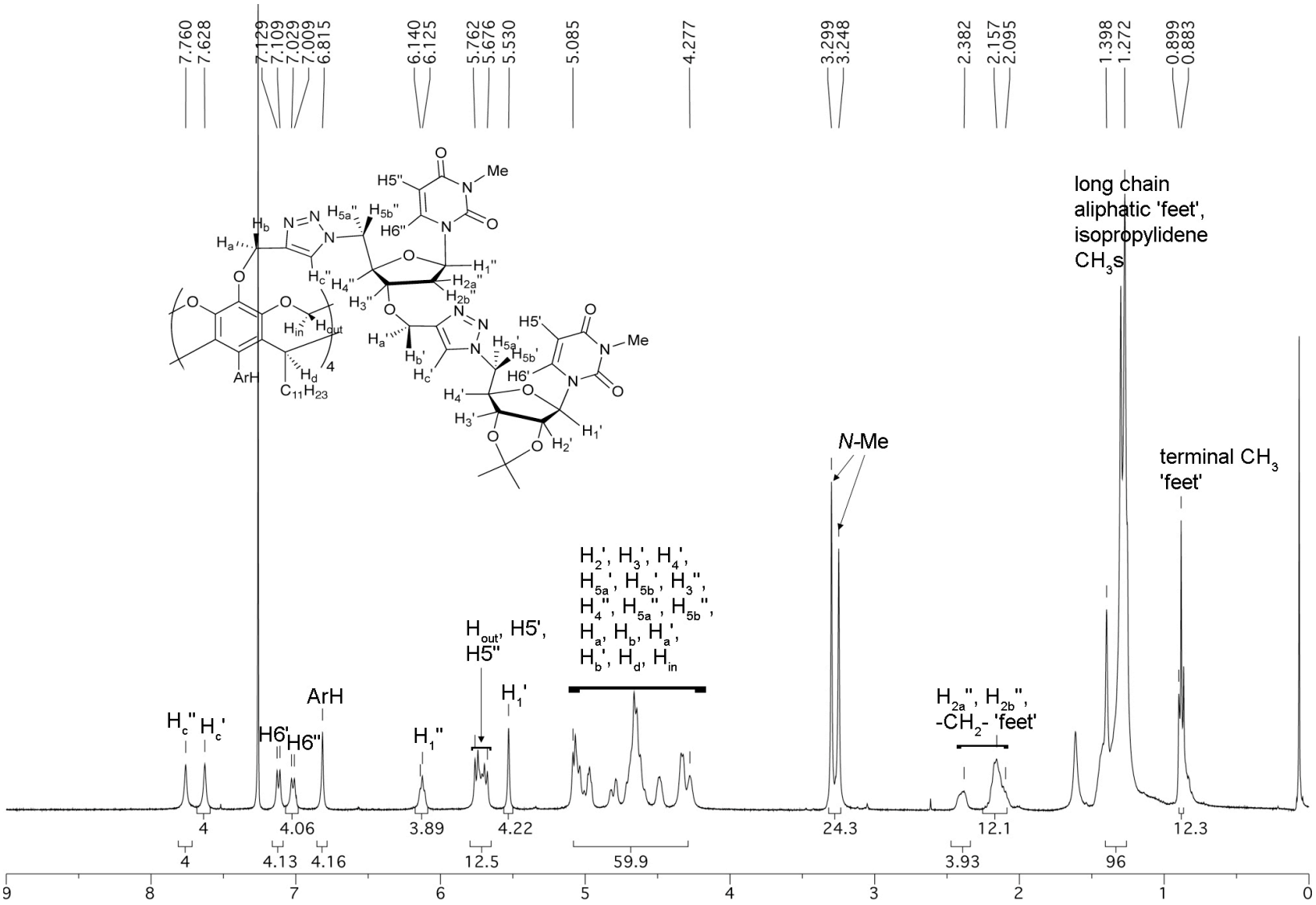


Figure 3.16 ^{1}H NMR spectrum of 27 in $CDCl_3$ at 25 °C.

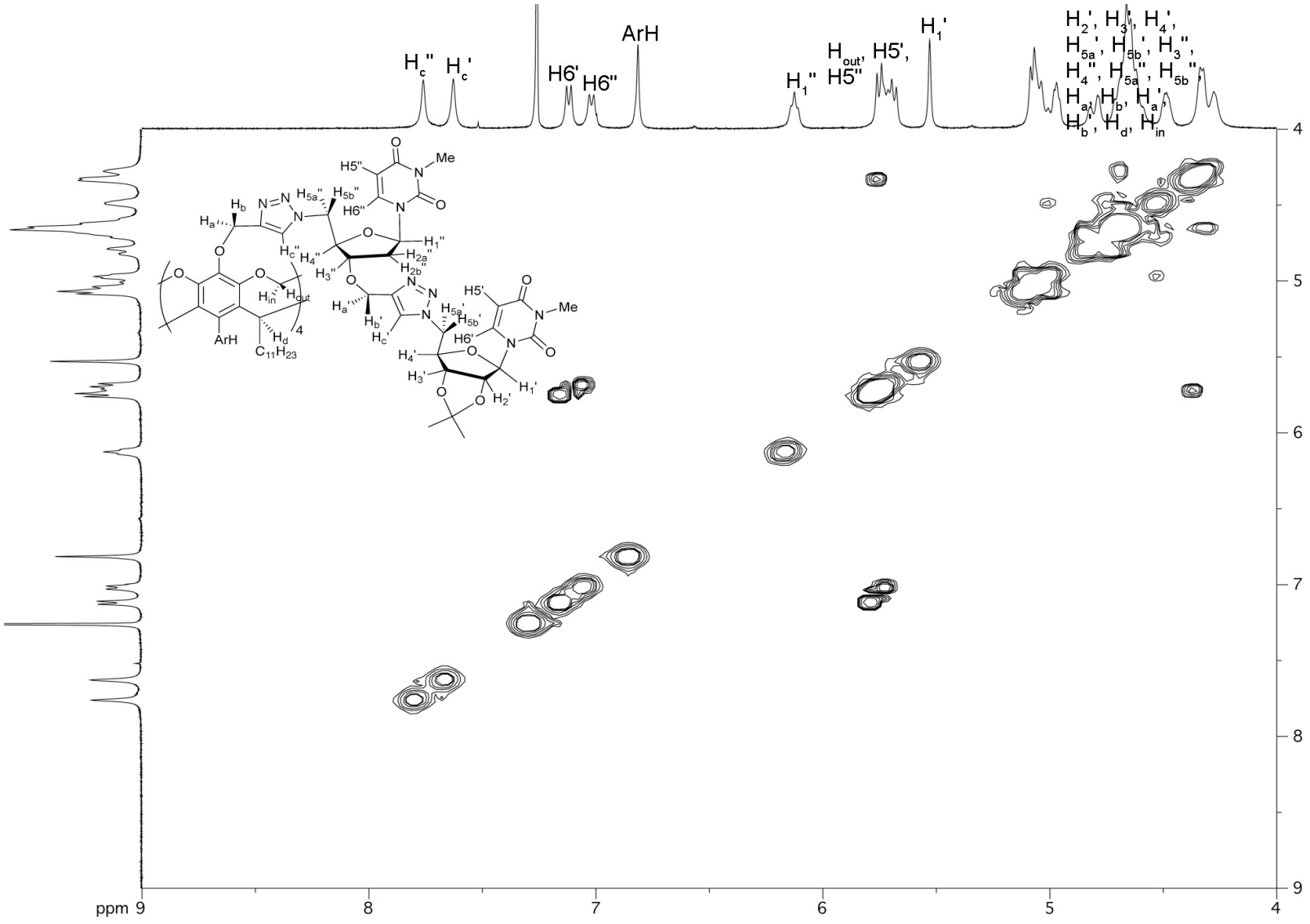


Figure 3.17 ^{1}H-^{1}H COSY spectrum of 27 in $CDCl_3$ at 25 ºC.

Chapter 4: Self-assembly of a Thymine Quartet and Quadruplex via an Organic Template

4.1 Synopsis

The previous chapters have described both uracil-based quartet and quadruplex constructs assembled from uridine units covalently coupled to a lipophilic cavitand. It was next sought to investigate the propensity of template-supported thymidine residues to adopt similar T-quartet and quadruplexes. Computational studies have indicated the decreased stability of T-quartets due to the additional steric bulk of the 5-Me group on thymine.[89] Nevertheless, Bare and Sherman recently reported a discrete T-quartet assembled at the foot of a hydrophilic cavitand in methanol.[130] This provided the impetus for the construction and characterization of its lipophilic counterpart, as well as the quadruplex.

This chapter documents the synthesis and characterization of conjugates **33** and **34**, which are the thymine analogs of **1** and **18** respectively (Figure 4.1). Comprehensive spectral studies were performed on the conjugates for complete ^{1}H NMR signal assignment and to probe for the assembly of a quartet and quadruplex.

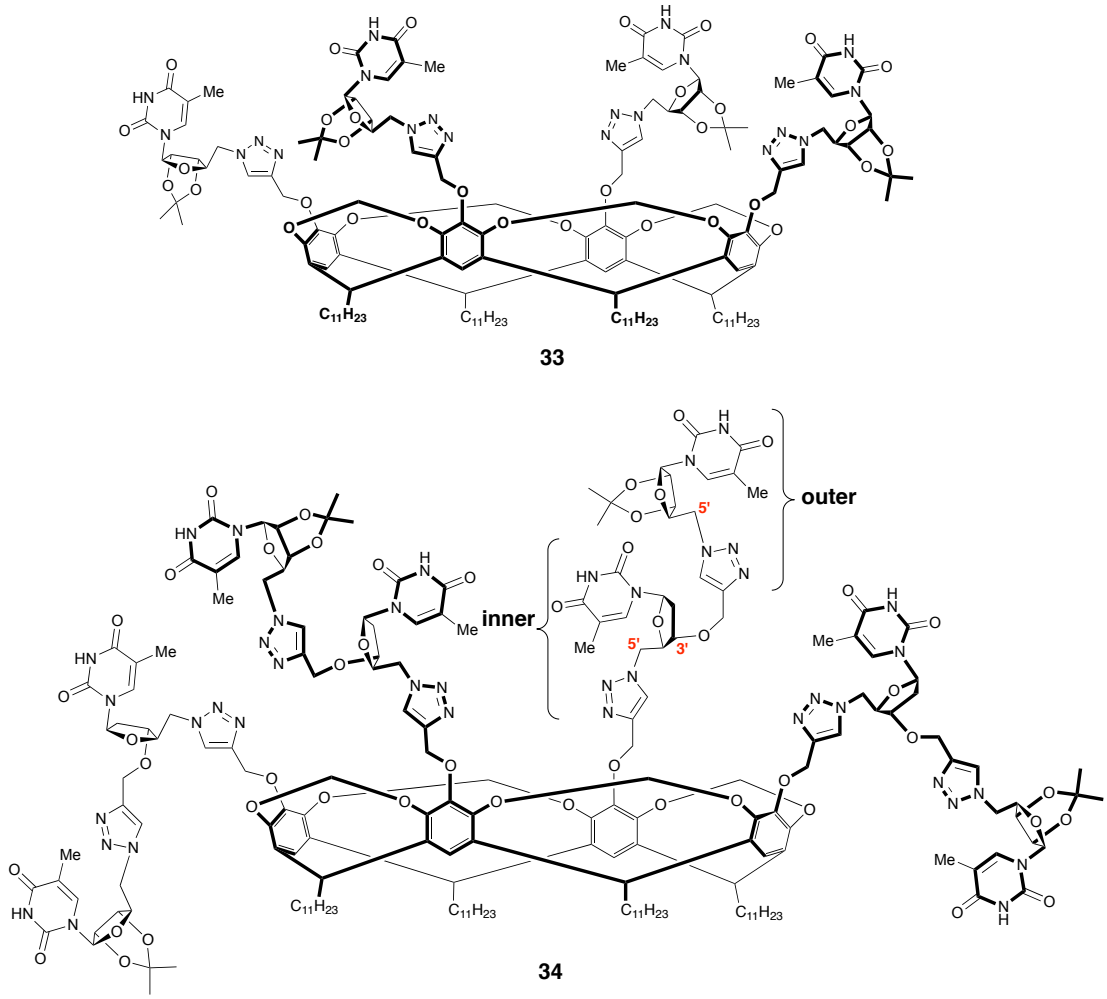


Figure 4.1 Side-on representations of conjugates 33 and 34. Outer and inner thymidine residues for 34 are labeled, as well as the 5' and 3' ribose ring positions showing the 5' – 3' outer – inner directional sense.

4.2 CPK modeling studies

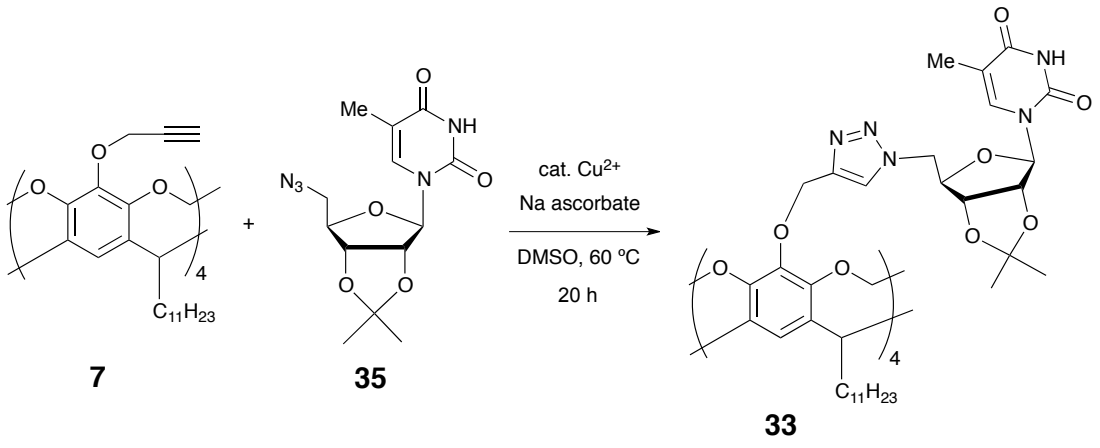
In contrast to the U-quartets of **1** and **18**, CPK models of **33** and **34** with the T-quartets set in-plane show the steric conflict between the projecting 5-Me group of each thymine base with the O2 atom of its neighbor. This leads to two possible structural modifications as proposed by theoretical models: 1) the tilting of the thymine rings in a propeller-twisted configuration to position the methyl groups off-plane or 2) an increase of the interbase distance leading to the

lengthening of the hydrogen bonds. The former allows for hydrogen bond lengths close to that of a U-quartet, while the latter maintains a planar but more weakly associated T-quartet. The CPK models do not provide information as to which state might be favored, but reflect theoretical expectations of unstable T-quartets relative to U-quartets.

4.3 Synthetic strategy

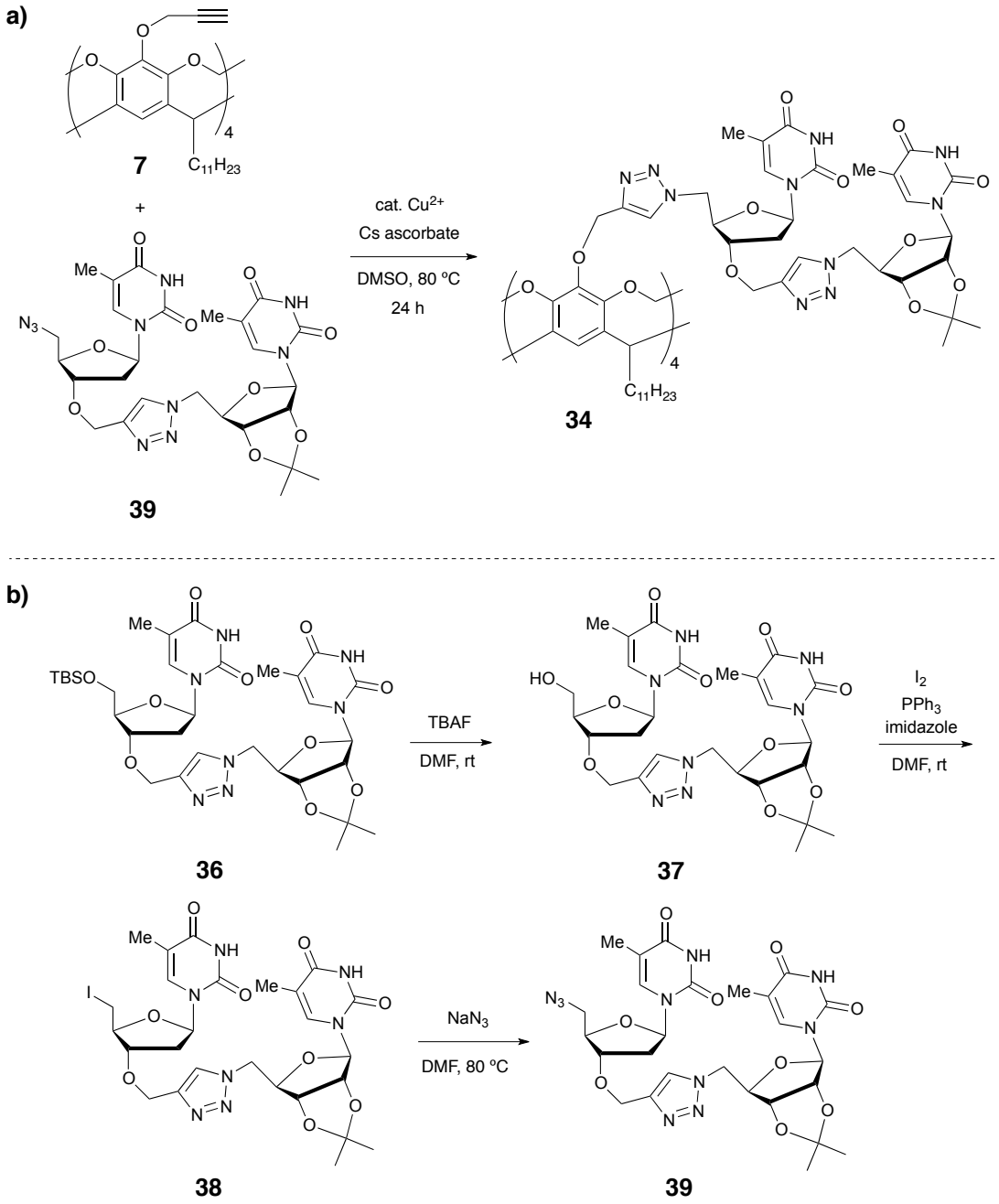
4.3.1 Synthesis of conjugates 33 and 34

The synthesis of **33** involved the Cu-catalyzed "click" coupling between cavitand **7** and azide **35** (Scheme 4.1). Molecular weight confirmation was provided by MALDI-TOF analysis. The conjugate is soluble in chloroform and methanol. **35** itself was prepared from 5-methyluridine via a similar procedure to that of **10**.



Scheme 4.1 Cu-catalyzed "click" reaction between 7 and 35 to access conjugate 33.

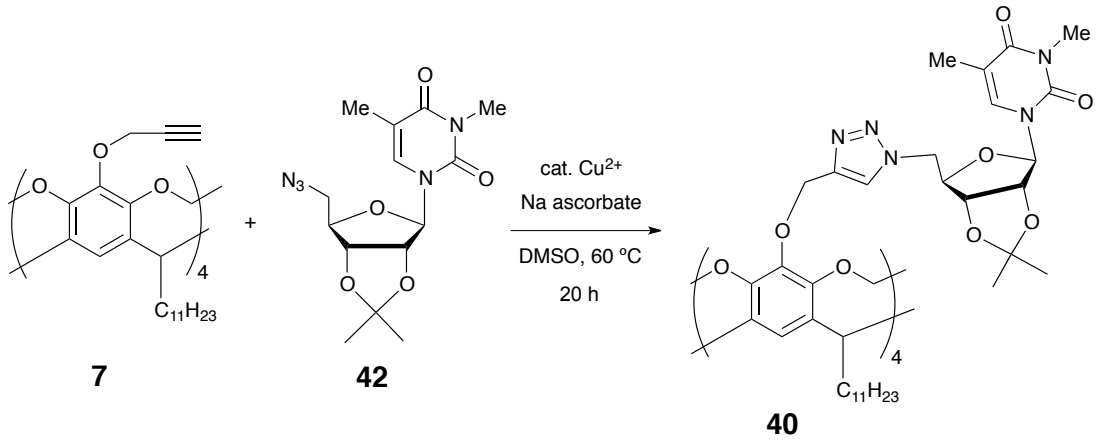
The preparation of conjugate **34** entailed the Cu-catalyzed "click" reaction between cavitand **7** and dinucleoside **39** (Scheme 4.2a), which was prepared via a similar procedure to that of **25** with the exception of the TBAF deprotection of silyl ether **36** (Scheme 4.2b). DMF was used as the solvent in this step due to the aggregation of **36** in THF to form an insoluble gel. This was also observed for **38** and **39** in chloroform. Thus, purification of these compounds could not be performed with standard silica gel column chromatography, and instead were re-precipitated and dried. These crude compounds were found to be sufficiently pure by NMR and were used directly in subsequent steps. The molecular weight of **34** was confirmed by MALDI-TOF analysis and it is freely soluble in chloroform and DMSO.



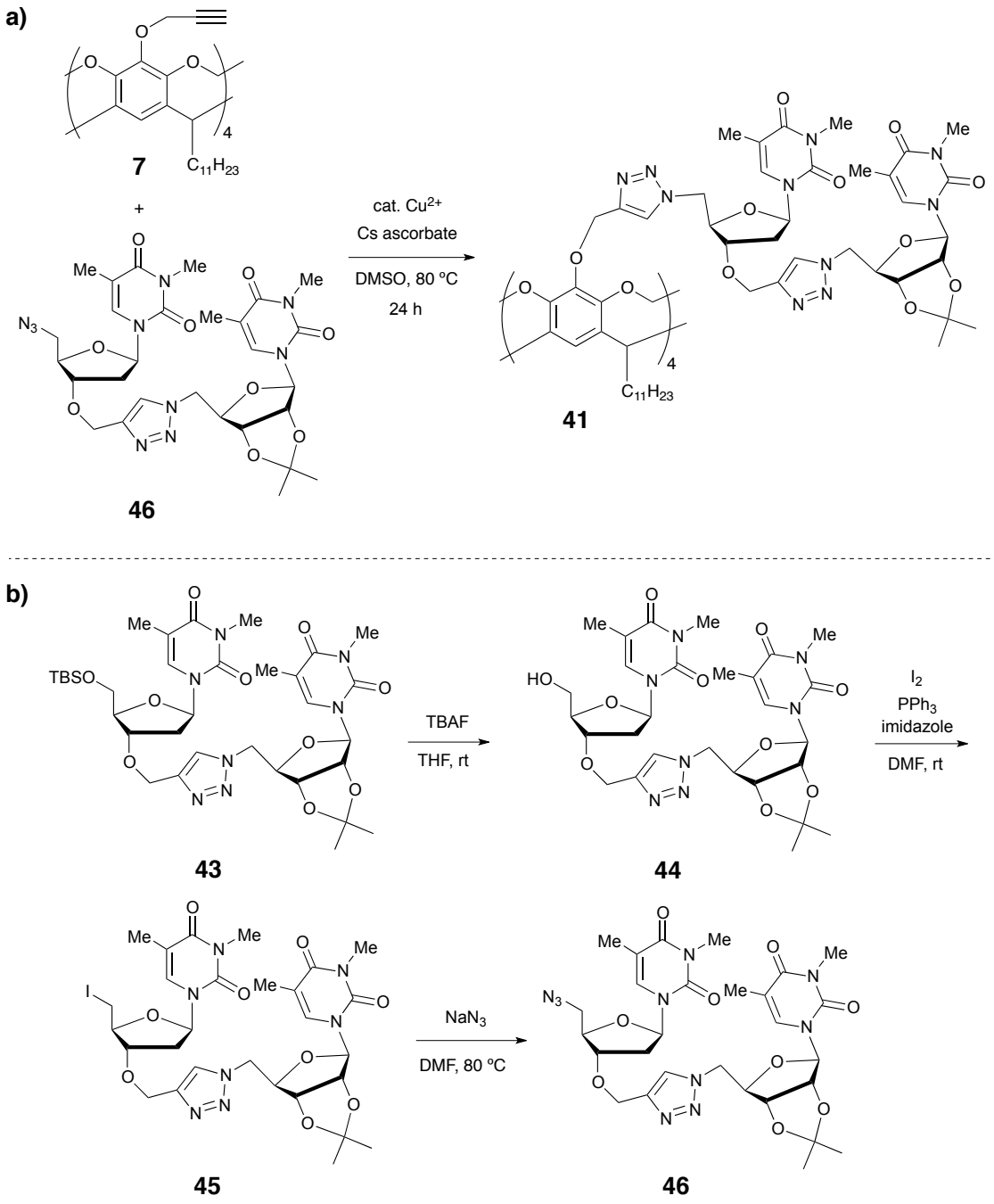
Scheme 4.2 a) Cu-catalyzed "click" reaction between 7 and 39 to afford conjugate 34. b) Synthesis of dinucleoside 39.

4.3.2 Synthesis of *N*-methylated analogs

Conjugates **40** and **41** with fully methylated imino positions were prepared via similar methods as those of uracil-based **17** and **27** respectively (Schemes 4.3 and 4.4). Molecular weight confirmation was provided by MALDI-TOF analysis. **40** exhibits good solubility in chloroform and methanol, and **41** in chloroform.



Scheme 4.3 Synthesis of conjugate 40.



Scheme 4.4 a) Cu-catalyzed "click" reaction between 7 and 46 to afford conjugate 41. b) Synthesis of dinucleoside 46.

4.4 NMR characterization studies

4.4.1 ^{1}H signal assignments

The ^{1}H NMR spectrum of **33** in $CDCl_3$ at room temperature is broadened with respect to **1**, likely due to conformational exchange. The imino, H_c, H6, ArH and 5-Me signals were assigned in $CDCl_3$ and are listed in Table 4.1. Full signal assignment was completed in DMSO-d_6, and the ^{1}H and ^{1}H-^{1}H COSY spectra are appended in the Experimental Section. Figure 4.3a shows the condensed structure of **33** with all protons labeled. *N*-methylated analog **40** was likewise fully assigned in DMSO-d_6 and the H_c, ArH, *N*-Me and 5-Me signals were assigned in $CDCl_3$.

Severe signal broadening is observed for **34** in $CDCl_3$ at room temperature, similar to **18**. In this case, the imino signals are resolved into two distinguishable peaks, whereas in **18** they are coincided (Figure 4.2). Full ^{1}H signal assignment of **34** was carried out in DMSO-d_6, and the ^{1}H and ^{1}H-^{1}H COSY spectra are provided in the Experimental Section. Figure 4.3b shows the condensed structure of **34** with all protons labeled. *N*-methylated analog **41** was likewise fully assigned in DMSO-d_6 and the H_c, ArH, *N*-Me and 5-Me signals were assigned in $CDCl_3$.

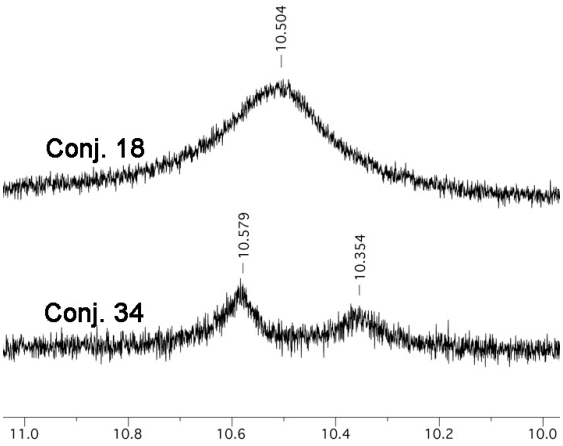


Figure 4.2 Portions of 400 MHz ^{1}H spectra of 18 and 34 in $CDCl_3$ at 25 ºC showing the imino proton signals.

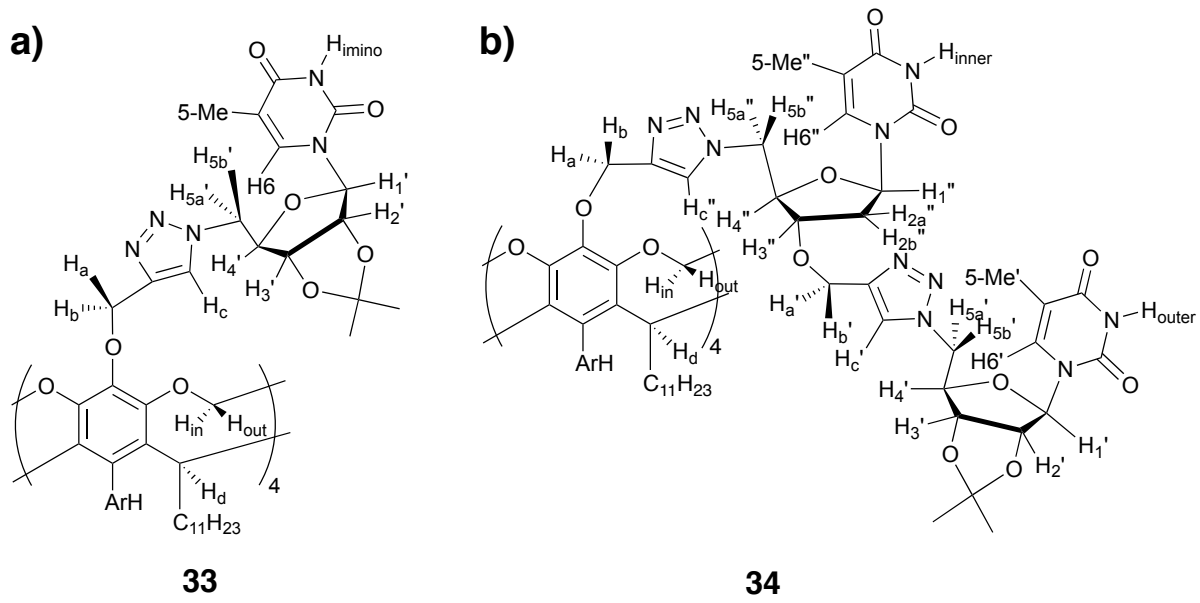


Figure 4.3 Condensed structures of a) 33 and b) 34 with all protons labeled.

4.4.2 Regiochemical analysis by NOESY

As with the previously reported conjugates, the 1,4-regiochemistry of all triazole moieties in **33**, **34**, **40** and **41** was confirmed via NOE correlations between the respective H_c and flanking methylene linker protons (Figure 4.4).

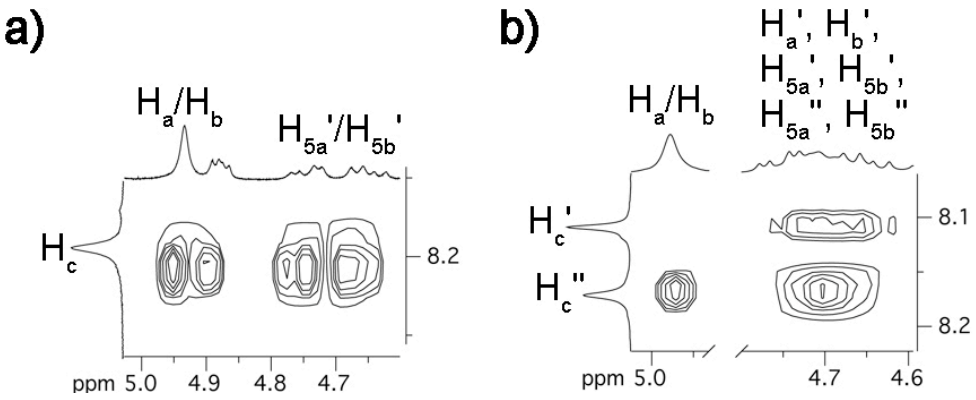


Figure 4.4 Portions of 400 MHz NOE spectra in DMSO-d_6 at 25 ºC showing appropriate crosspeaks signifying 1,4-regiochemistry at both triazole rings for a) 33 and b) 34. Conjugates 40 and 41 displayed similar crosspeaks and are therefore not shown.

4.4.3 Summary of ^{1}H signal assignments

Tables 4.1 through 4.4 summarize the ^{1}H signal assignments for **33**, **34**, **40** and **41** respectively.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
H_{imino}	10.87	11.43		
5-Me	1.92	1.72	H6	H6
H6	7.00	7.42	5-Me (long range)	5-Me
H_1'		5.74	H_2'*	
H_2'		5.03	H_1'*, H_3'	
H_3'		4.88	H_2', H_4'**	
H_4'		4.38	H_3'**, H_{5a}'/H_{5b}'	
H_{5a}'		4.74	H_4', H_{5b}'	
H_{5b}'		4.65	H_4', H_{5a}'	
H_a/H_b		4.93		H_c
H_c	7.70	8.18		H_a/H_b, H_{5a}'/H_{5b}'
H_d		4.55	-CH_2- "feet"	ArH
H_{in}		4.26	H_{out}	
H_{out}		5.90	H_{in}	
ArH	6.84	7.27		H_d
-CH_2- 'feet'		2.31	H_d	
long chain 'feet', isoprop. CH_3		1.22 – 1.44		
terminal CH_3		0.84		

* Expected but unobservable COSY signals.

** Weak COSY signals.

Table 4.1 ^{1}H signal assignments of 33 at 25 ºC.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
H_{outer}		11.45		
H_{inner}		11.31		
H_c'		8.11		H_a'/H_b', H_{5a}'/H_{5b}'
H_c"		8.17		H_a/H_b, H_{5a}"/H_{5b}"
H6'		7.47	5-Me'	5-Me'
H6"		7.36	5-Me"	5-Me"
5-Me'		1.74	H6'	H6'
5-Me"		1.77	H6"	H6"
ArH	6.81	7.27		H_d
H_1'		5.74	H_2'	
H_1"		6.10	H_{2a}"/H_{2b}"	
H_2'		5.07	H_1', H_3'*	
H_{2a}"/H_{2b}", -CH_2- 'feet'		2.16 – 2.29	H_{2b}"/H_{2a}"*	
H_3'		4.88	H_2', H_4'	
H_{out}		5.90	H_{in}	
H_a/H_b		4.97	H_b/H_a	
H_{5a}'/H_{5b}', H_{5a}"/H_{5b}", H_a'/H_b', H_d		4.53 – 4.77		
H_3", H_4', H_4", H_{in}		4.28 – 4.39		
long chain 'feet', isopropylidene CH_3		1.23 – 1.43		
terminal CH_3		0.84		

* Correlation indistinguishable due to signal overlapping in the region.

Table 4.2 ^{1}H signal assignments of 34 at 25 ºC.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.32	3.16		
5-Me	1.88	1.78	H6	H6
H6	7.02	7.51	5-Me	5-Me
H_1'		5.80	H_2'*	
H_2'		5.08	H_1'*, H_3'	
H_3', H_a/H_b		4.93 – 4.95	H_2', H_b/H_a	
H_4'		4.40	H_3'**, H_{5a}'/H_{5b}'	
H_{5a}'/H_{5b}'		4.65 – 4.79	H_4', H_{5b}'/H_{5a}'	
H_c	7.66	8.18		H_a/H_b, H_{5a}'/H_{5b}'
H_d		4.58	-CH_2- "feet"	ArH
H_{in}		4.26	H_{out}	
H_{out}		5.89	H_{in}	
ArH	6.79	7.27		H_d
-CH_2- 'feet'		2.30	H_d	
long chain 'feet', isoprop. CH_3		1.24 – 1.46		
terminal CH_3		0.85		

* Expected but unobservable COSY signals.

** Weak COSY signals.

Table 4.3 ^{1}H signal assignments of 40 at 25 °C.

Proton	δ_H $CDCl_3$ (ppm)	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.27	3.12		
N-Me	3.32	3.16		
H_c'	7.62	8.10		H_a'/H_b', H_{5a}'/H_{5b}'
H_c"	7.73	8.15		H_a/H_b, H_{5a}"/H_{5b}"
5-Me'	1.85	1.79	H6'	H6'
5-Me"	1.91	1.80	H6"	H6"
H6'	6.98	7.54	5-Me'	5-Me'
H6"	6.80	7.41	5-Me"	5-Me"
H_1'		5.77	H_2'	
H_1"		6.14	H_{2a}"/H_{2b}"	
H_2'		5.09	H_1'*, H_3'	
H_{2a}"/H_{2b}", -CH_2-'feet'		2.19 – 2.35	H_{2b}"/H_{2a}"*, H_d	
H_3', H_a/H_b		4.90 – 4.97	H_2', H_4'*, H_b/H_a	
H_3", H_4', H_4", H_{in}		4.25 – 4.41		
H_{5a}'/H_{5b}', H_{5a}"/H_{5b}", H_a'/H_b', H_d		4.51 – 4.79		
H_{out}		5.88	H_{in}	
ArH	6.78	7.26		H_d
long chain 'feet', isoprop. CH_3		1.22 – 1.44		
terminal CH_3		0.84		

* Correlation indistinguishable due to signal overlapping in the region.

Table 4.4 ^{1}H signal assignments of 41 at 25 ºC.

4.4.4 NMR solution structure elucidation

4.4.4.1 Conjugate 33

^{1}H NMR spectra of **33** in $CDCl_3$ recorded at 25, –20 and –40 ºC show a steady downfield shift of the imino proton as the temperature is decreased, indicating its increasingly deshielded, hydrogen bound state (Figure 4.5a). At –20 and –40 ºC, a second, weaker imino signal of similar chemical shift is observed approximately 0.1 ppm downfield and upfield respectively of the more intense parent signals. NOESY performed at room temperature and –20 ºC shows no correlation between the imino and 5-Me protons, indicating the absence of quartets. The two imino signals observed at –20 ºC might thus arise from an observable exchange between major and minor non-quartet hydrogen-bound states whose conformations are unknown. At –40 ºC, the crosspeak is apparent for the imino signal at 11.8 ppm (Figure 4.5b), and not for the other, suggesting a major T-quartet species in exchange with a minor non-quartet state possibly identical to that found at –20 ºC with the same chemical shift (11.7 ppm). In accordance with the theoretical model,[89] this T-quartet is proposed to adopt the more energetically favorable propeller-twist arrangement. In contrast, no correlations between the *N*-Me and 5-Me positions of **40** were observed at –40 ºC in $CDCl_3$, confirming the central importance of the imino proton in the T-quartet assembly.

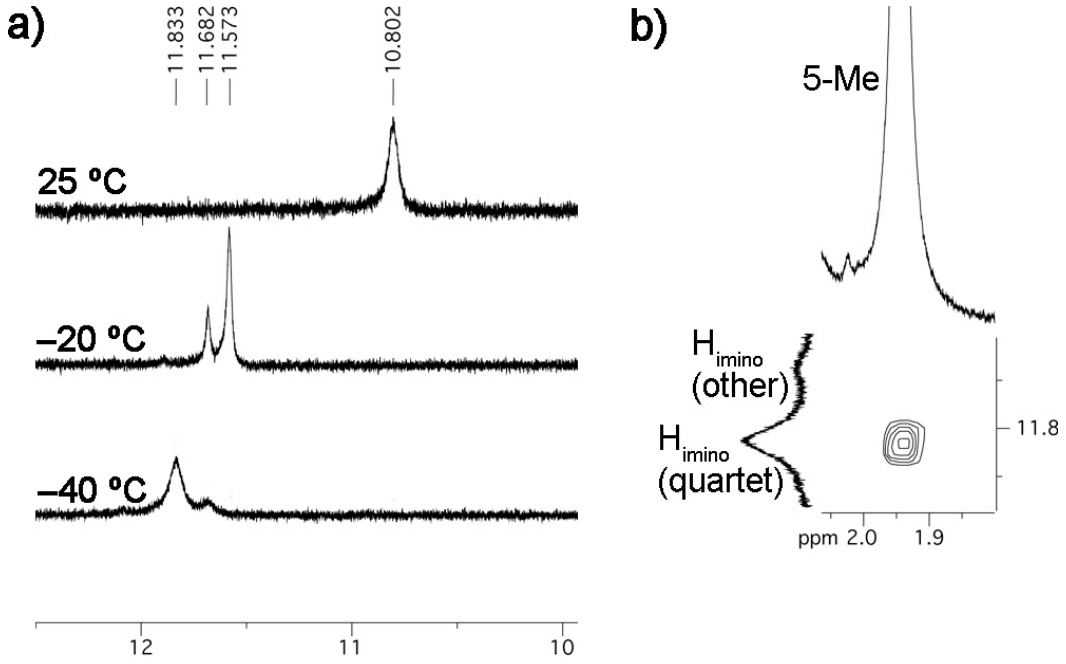


Figure 4.5 a) Portions of 400 MHz variable-temperature ^{1}H spectra of 33 in $CDCl_3$. Hydrogen-bound species not equivalent to, but approaching an ideal T-quartet are proposed to give rise to the signals at –20 ºC. Major species at –40 ºC is attributed to ideal T-quartet with minor non-ideal species in existence. b) Portion of 400 MHz NOE spectrum of 33 in $CDCl_3$ at –40 ºC showing crosspeak between the imino proton and 5-Me indicative of a T-quartet.

Proximity between the thymine bases and underlying triazole linkers of **33** was indicated by observed NOE correlations between H_c and both H6 and 5-Me (Figure 4.6a) positions at –40 ºC, suggesting the stacking of these ring systems (Figure 4.6b). These interactions are neither observed at 25 ºC nor at –20 ºC confirming the lack of structure at these temperatures.

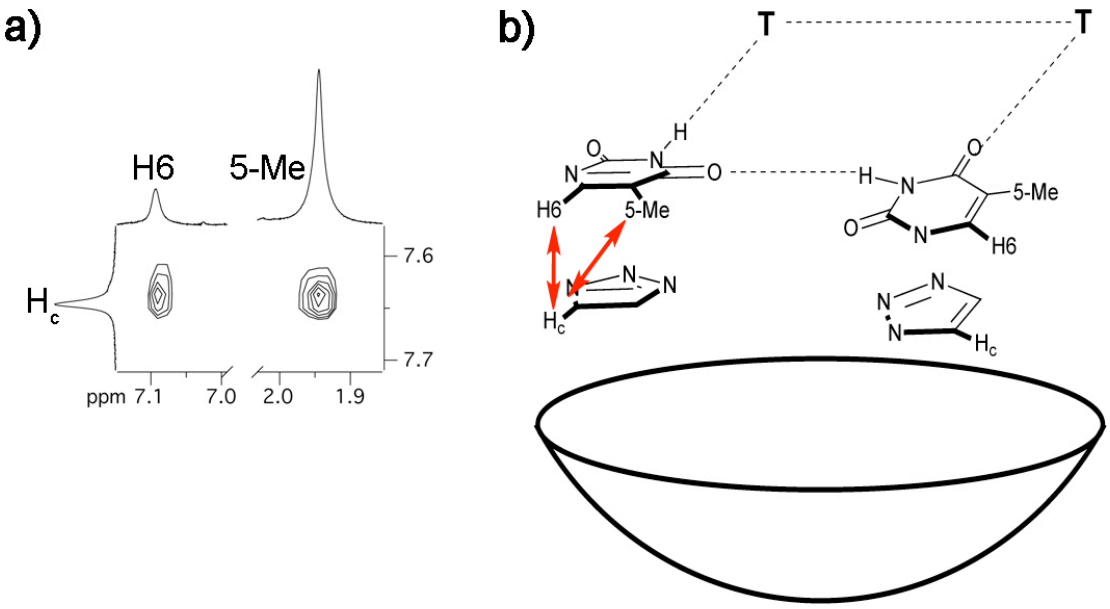


Figure 4.6 a) Portion of 400 MHz NOE spectrum of 33 in $CDCl_3$ at –40 ºC showing correlations between H_c, H6 and 5-Me. b) Illustration of proposed twisted T-quartet of 33 illustrating stacked thymine and triazole rings with NOE interactions shown.

The *syn* glycosidic conformation of the thymine bases in **33** is evidenced by the characteristic NOE correlation between H6 and H_1' (Figure 4.7).

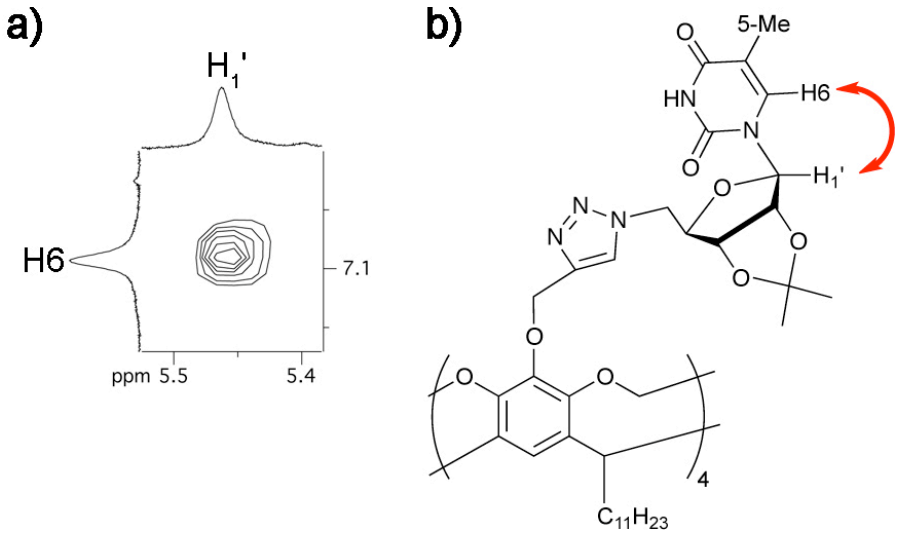


Figure 4.7 a) Portion of 400 MHz NOE spectrum of 33 in $CDCl_3$ at -40 ºC showing the H6/H_1' crosspeak indicative of the *syn* glycosidic bond. b) The *syn* conformation of 33 with H6/H_1' interaction shown.

The oligomeric state of **33** in $CDCl_3$ at room temperature with respect to **40** was determined by DOSY NMR. Table 4.5 lists the measured diffusion constants of H_c for both compounds that are in close agreement, indicating their identical oligomeric states. Since **40** is unable to form hydrogen bonds, it is assumed to be unimolecular under the present conditions, and therefore **33** is also unimolecular.

Compound	*D* (x10^{-10} m^2 s^{-1})	Ratio
33	3.69 ± 0.10	0.95
40	3.46 ± 0.10	

Table 4.5 H_c diffusion constants (*D*) of 33 and 40 in $CDCl_3$ at 25 ºC (2.7 mM).

Attempts to conduct DOSY measurements at –40 ºC were unsuccessful due to instrument limitations. Temperature stability is crucial in DOSY experiments, as any fluctuation generates convection currents in the solvent, especially one of low viscosity such as $CDCl_3$, resulting in inaccurate measurements. Maintenance of a stable temperature as low as –40 ºC was not possible with the current instrumentation, and intermittent temperature spikes of up to 1 ºC were observed. Reliable readings thus could not be obtained. This problem could not be circumvented with the use of smaller diameter NMR tubes. It is noted, however, that the ^{1}H NMR spectrum of **33** in $CDCl_3$ at –40 ºC shows no evidence of aggregation as the signals remain well resolved.

4.4.4.2 Conjugate 34

The NOE spectrum of **34** in DMSO-d_6 at room temperature shows crosspeaks of different intensities between both sets of imino protons and the respective 5-Me positions, consistent with a quadruplex assembly composed of a more tightly bound inner quartet (H_{inner} = 11.32 ppm, 5-Me" = 1.77 ppm) capped by an outer quartet (H_{outer} = 11.45 ppm, 5-Me' = 1.75 ppm) similar to the U-quadruplex system reported in **18** (Figure 4.8a). It is proposed that the inner T-quartet assumes the propeller-twisted configuration, allowing for shorter and thus stronger hydrogen bonds, while the outer T-quartet is planar giving rise to longer and weaker hydrogen bonds and hence a more weakly associated system. NOE connections between H_c" and H6", as well as H_c' to both H6' and H6", position all four ring systems (both triazole and both thymine) in close proximity indicating intercalation of the triazole and thymine rings in a fully stacked quadruplex (Figure 4.8b). Thus, ring stacking via π-π interactions is expected to be a major contributor to the stability of this scaffold since thymine systems are unstable with respect

to uracil. This was investigated and is discussed in Section 4.4.4.3. A visualization of the proposed assembly is presented in Figure 4.9c. NOE interactions are also displayed therein.

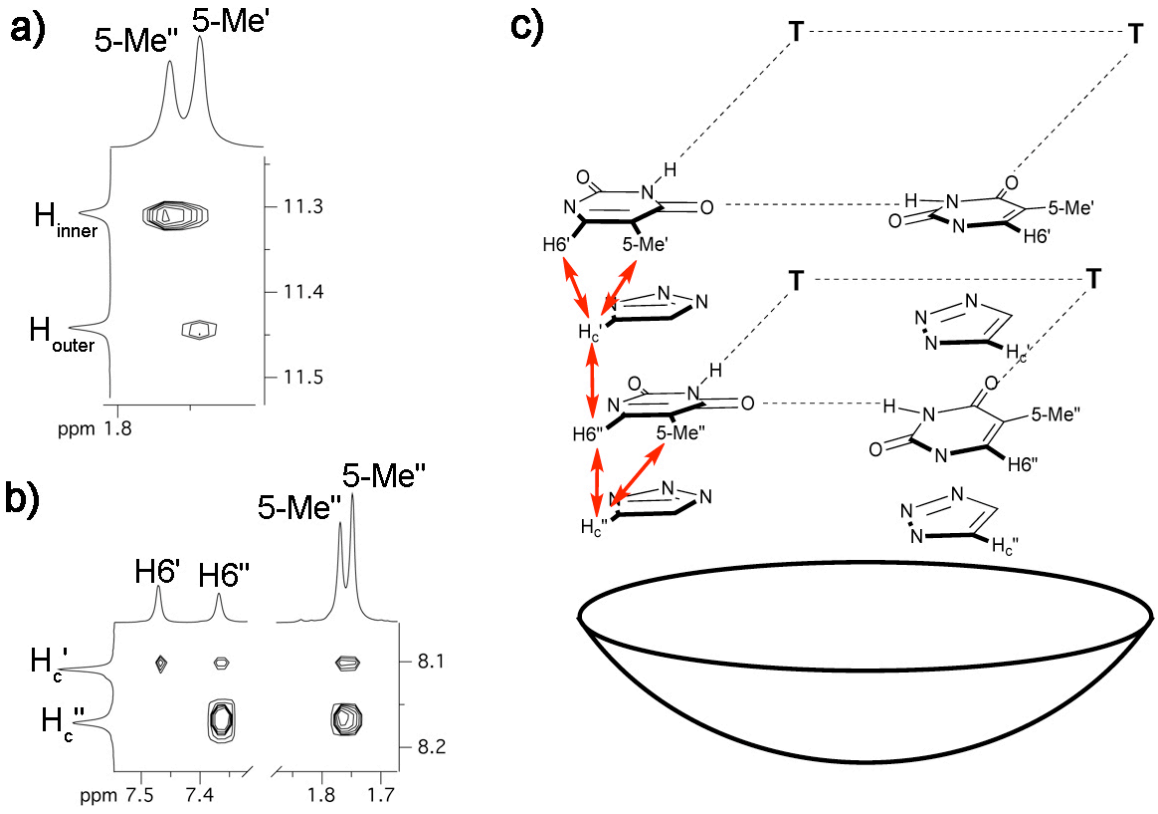


Figure 4.8 Portions of 400 MHz NOE spectrum of 34 in DMSO-d_6 at 25 ºC showing a) Imino/5-Me crosspeaks indicative of quadruplex assembly and b) H_c/H5/5-Me crosspeaks indicative of a fully intercalated scaffold. c) Illustration of the quadruplex assembly in 34 with planar outer quartet and propeller-twisted inner quartet. NOE connections are labeled.

NOESY performed on **41** in DMSO-d_6 at 25 ºC revealed no correlations between either sets of *N*-Me and 5-Me signals, confirming the absence of quartets as expected. Likewise, no correlations were found between the imino and 5-Me protons of dinucleoside **39**, although water-mediated exchange between the imino protons is apparent from their mutual crosspeaks as

well as to residual water in the NMR solvent (Figure 4.9), a process also observed in the uracil analog **25**. Solvent exclusion by the closed, hydrogen bound quadruplex system in **34** precludes this process.

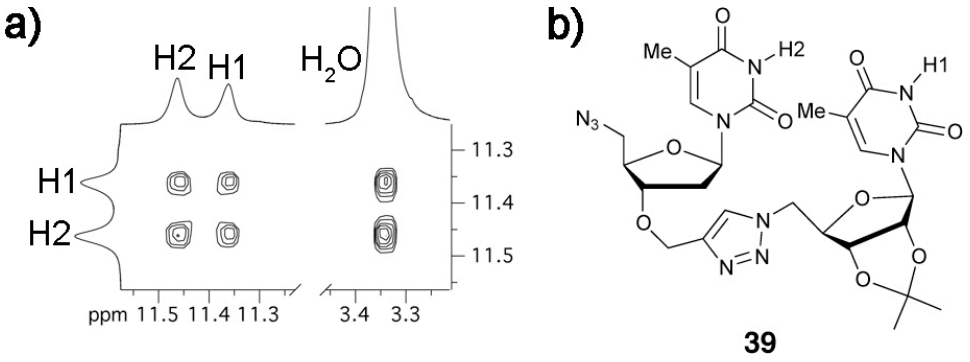


Figure 4.9 a) Portion of 400 MHz NOE spectrum of 39 in DMSO-d_6 at 25 ºC showing mutual imino crosspeaks and to H_2O. b) Imino protons labeled in the structure of 39.

Syn glycosidic bonds are present in both the ribose rings of **34** as represented by NOE crosspeaks between H6' and H_1" as well as H6" and H_1" (Figure 4.10).

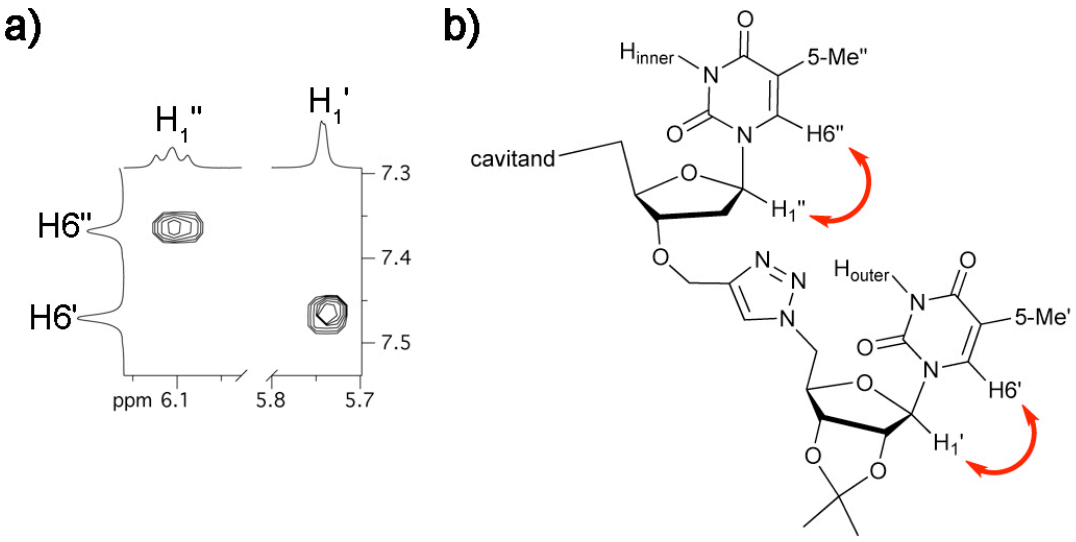


Figure 4.10 a) Portion of 400 MHz NOE spectrum of 34 in DMSO-d_6 at 25 ºC showing crosspeaks between the H6 and H_1 protons of both nucleosides. b) NOE correlations in the *syn* configuration of 34.

DOSY measurements of **34** and **41** revealed the unimolecularity of both compounds in DMSO-d_6 at room temperature, as the diffusion constant for the observed H_c" protons are in close agreement to within 4%. In $CDCl_3$, the ArH protons were observed instead due to the overlap of H_c' and H_c" for **34**. The diffusion constant for **34** in $CDCl_3$ was found to be 81% that of **41**, revealing its bimolecular state and confirming the earlier hypothesis of the aggregation of **34** in $CDCl_3$. Table 4.6 summarizes all the measured diffusion constants for both conjugates.

Solvent	Signal	34	41	Ratio
$CDCl_3$	ArH	2.66 ± 0.10	3.29 ± 0.10	0.81
DMSO-d_6	H_c"	0.68 ± 0.10	0.71 ± 0.10	0.96

Table 4.6 Diffusion constants (*D*) of 34 and 41 in $CDCl_3$ and DMSO-d_6 at 25 ºC (2.7 mM).

4.4.4.3 Further investigations on the quadruplex ring stacking effect

To obtain further insight on the effect of ring stacking on the quadruplex assembly of **34**, hemimethylated conjugate **47** was designed in which only the outer thymine imino positions are methylated (Figure 4.11).

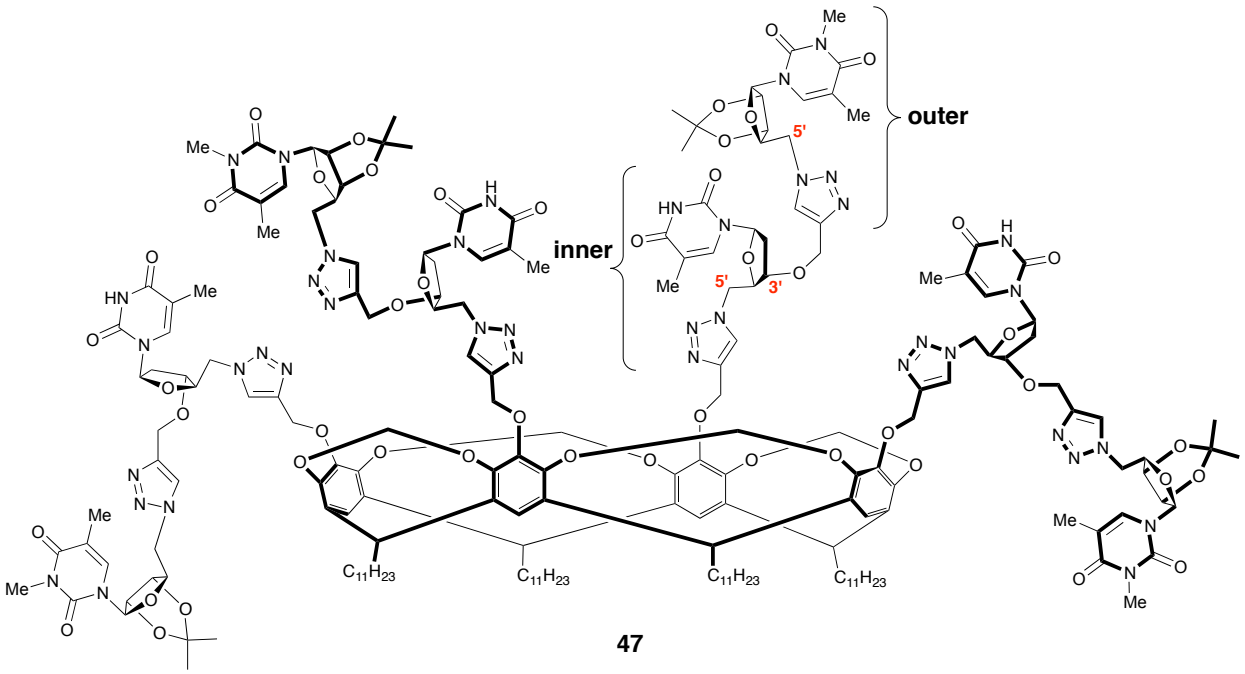
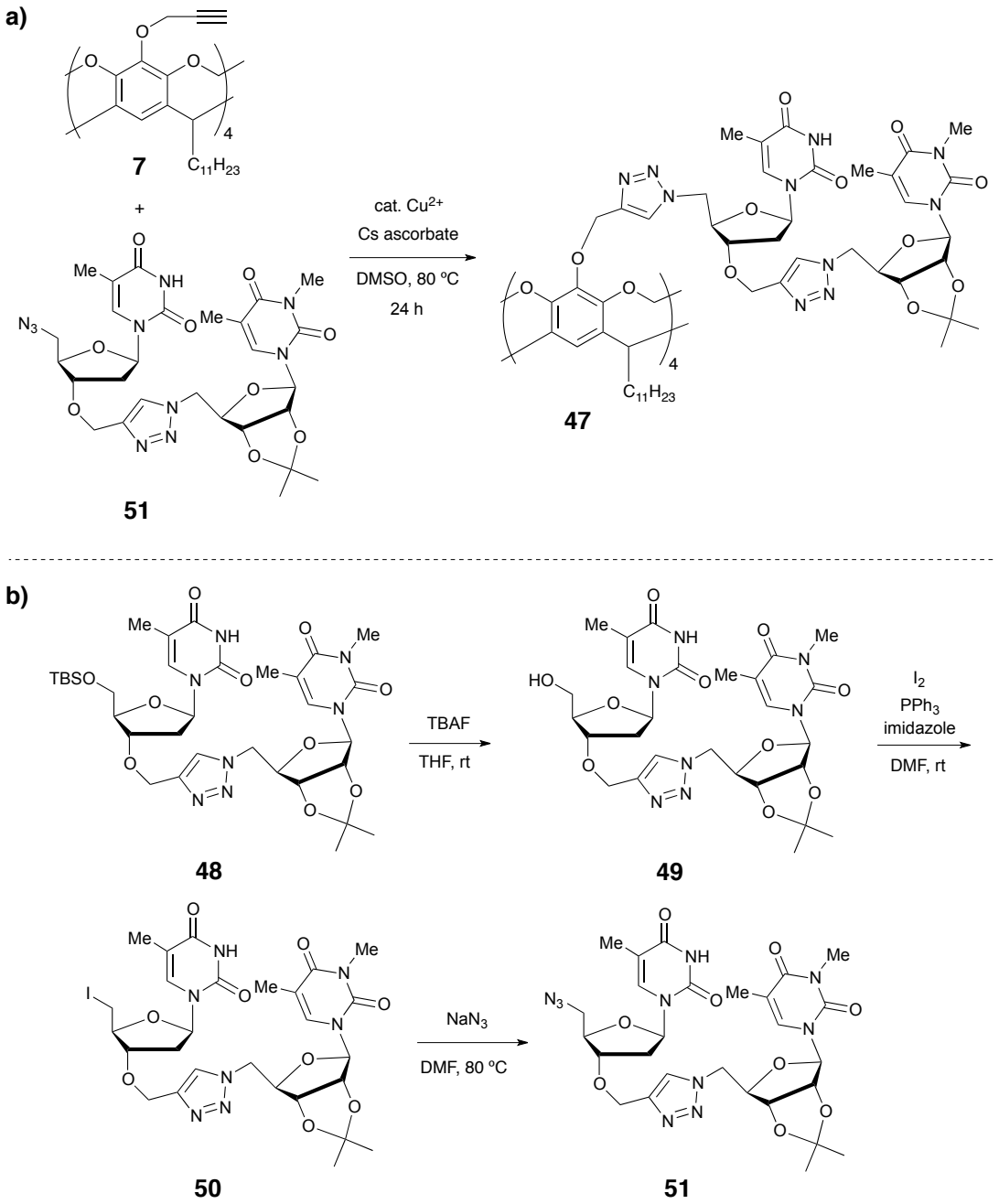


Figure 4.11 Side-on view of conjugate 47.

47 was accessed by the Cu-catalyzed "click" reaction between cavitand **7** and dinucleoside **51** (Scheme 4.5a), prepared with similar chemistry to that of **39** and **46** (Scheme 4.5b).



Scheme 4.5 a) Cu-catalyzed "click" reaction between 7 and 51 to afford conjugate 47. b) Synthesis of dinucleoside 51.

Assignment of the ^{1}H NMR signals was completed in DMSO-d_6 with the use of ^{1}H-^{1}H COSY and NOESY techniques as with all other conjugates described thus far in this work. 1,4-

regiochemistry at both triazole rings and all-*syn* glycosidic bonds were likewise confirmed with the observation of the appropriate NOE correlations (Figure 4.12). Table 4.7 lists all the signal assignments, and the ^{1}H and ^{1}H-^{1}H COSY spectra are provided in the Experimental Section.

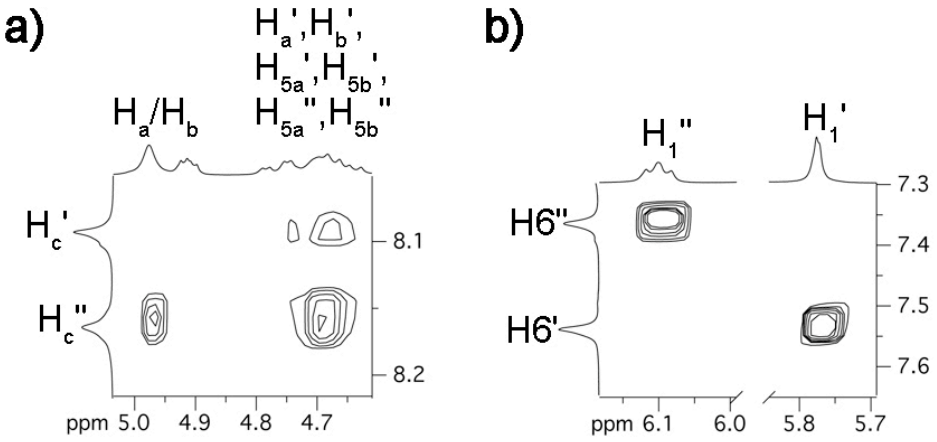


Figure 4.12 Portions of 400 MHz NOE spectrum of 47 in DMSO-d_6 at 25 ºC showing a) crosspeaks between triazole protons and the respective flanking methylene linkers indicative of uniform 1,4-regiochemistry and b) crosspeaks between H6 and H_1 protons of both thymidine residues indicative of all-*syn* glycosidic bond angles.

Proton	δ_H DMSO-d_6 (ppm)	COSY correlation	NOESY correlation
N-Me	3.16		
H_{inner}	11.31		
H_c'	8.09		H_a'/H_b', H_{5a}'/H_{5b}'
H_c"	8.16		H_a/H_b, H_{5a}"/H_{5b}"
H6'	7.54	5-Me'	5-Me'
H6"	7.36	5-Me"	5-Me"
5-Me'	1.80	H6'	H6'
5-Me"	1.77	H6"	H6"
ArH	7.27		H_d
H_1'	5.78	H_2'	
H_1"	6.10	H_{2a}"/H_{2b}"	
H_2'	5.09	H_1', H_3'*	
H_{2a}"/H_{2b}", -CH_2- 'feet'	2.16 – 2.28	H_{2b}"/H_{2a}"*	
H_3'	4.91	H_2', H_4'	
H_{out}	5.90	H_{in}	
H_a/H_b	4.98	H_b/H_a	
H_{5a}'/H_{5b}', H_{5a}"/H_{5b}", H_a'/H_b', H_d	4.55 – 4.79		
H_3", H_4', H_4", H_{in}	4.27 – 4.41		
long chain 'feet', isopropylidene CH_3	1.23 – 1.43		
terminal CH_3	0.85		

* Correlation indistinguishable due to signal overlapping in the region.

Table 4.7 ^{1}H signal assignments of 47 in DMSO-d_6 at 25 ºC.

The assembly of the inner T-quartet of **47** in DMSO-d_6 is indicated by an NOE correlation observed between H_{inner} and 5-Me" (Figure 4.13a). No crosspeak was observed between the *N*-Me and 5-Me' groups, ruling out the presence of an outer quartet as expected. Evidence of ring stacking was provided by H_c"/H6" and H_c'/H6' correlations (Figure 4.13b). Unlike **34**, the H_c'/H6" connection was not observed, but a H_c'/5-Me" crosspeak is visible instead to indicate proximity between the outer triazole and inner thymine. This could be due to a conformational distortion of the outer *N*-methylated thymine layer as a result of additional steric bulk imparted by the *N*-Me groups (Figure 4.13c), thereby rotating the triazole rings to position H_c' close to 5-Me" instead of H6". No evidence of a quartet assembly was observed in the NOE spectrum of **33** (lacking the outer layer) in DMSO-d_6, unmasking the crucial role of ring stacking in stabilizing the inner quartet of **47** and the overall quadruplex of **34**. π-π interactions are believed to be the driving force of this stacking. Of course, the contribution of the outer U- and T-quartets of **18** and **34** respectively towards overall U- and T-quadruplex stability is not discounted, and instead is believed to be substantial. Rather, the synergy of hydrogen bonding and π-π interactions in the stabilization of these otherwise unstable motifs has been demonstrated.

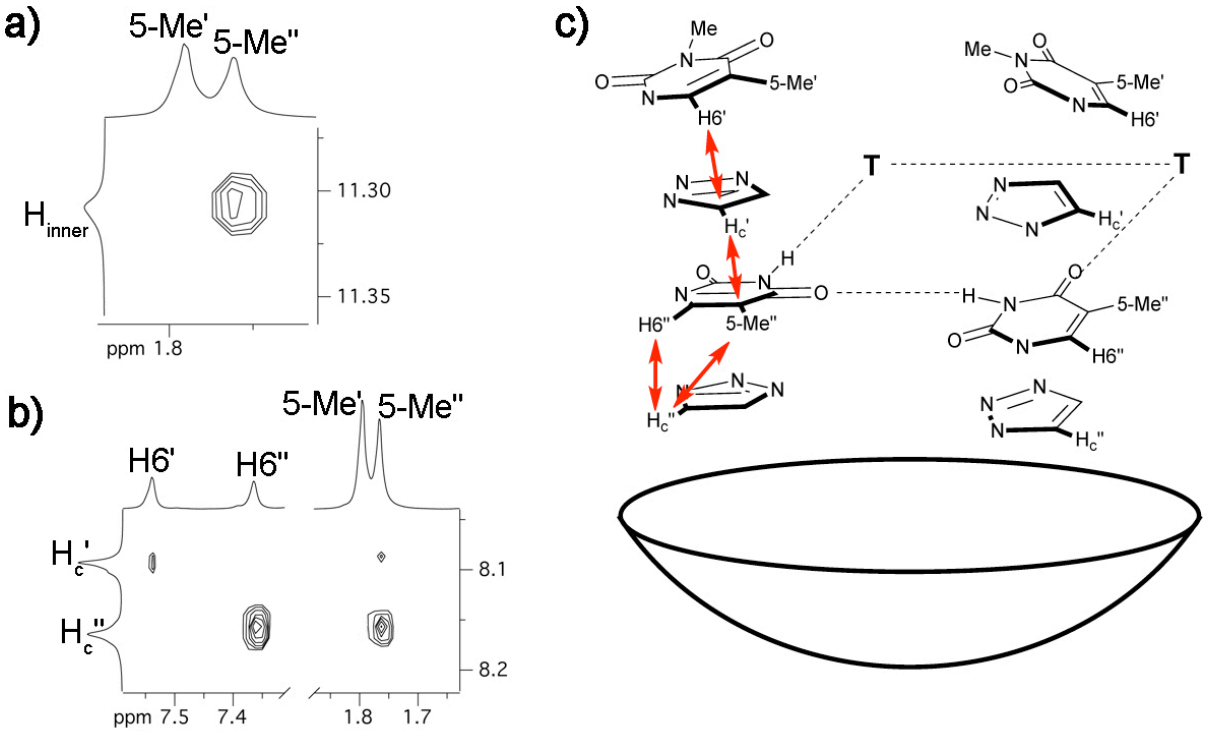


Figure 4.13 Portions of 400 MHz NOE spectrum of 47 in DMSO-d_6 at 25 ºC showing a) H_{inner}/5-Me" crosspeak indicative of an inner T-quartet assembly and b) crosspeaks implying ring stacking. c) Illustration of the proposed assembly in 47 with a distorted outer layer. NOE correlations are labeled with red arrows.

DOSY measurements confirmed that **47** remains unimolecular in DMSO-d_6 with the diffusion constant of H_c within 1% of **41** (Table 4.8).

Compound	D (x10^{-10} m^2 s^{-1})	Ratio
47	0.70 ± 0.10	0.99
41	0.71 ± 0.10	

Table 4.8 H_c" diffusion constants (D) of 34 and 41 in $CDCl_3$ and DMSO-d_6 at 25 ºC (2.7 mM).

4.4.5 CD spectroscopic studies of 33 and 40

The CD spectra of **33** and **40** in both chloroform and methanol at room temperature are indistinguishable with λ_{max} values at approximately 260 nm, indicating identical structural states and substantiating the NMR evidence of the non-existence of a quartet in **33** under these conditions (Figure 4.14). The spectrum of **1** in chloroform is provided for reference. Treatment of **33** with strontium picrate $Sr(pic)_2$ yielded no evidence of metal-induced quartet assembly, with the resultant spectrum indistinguishable from the others. These results further demonstrate the instability of the T-quartet in the solution state under ambient conditions, even in the presence of metal cations.

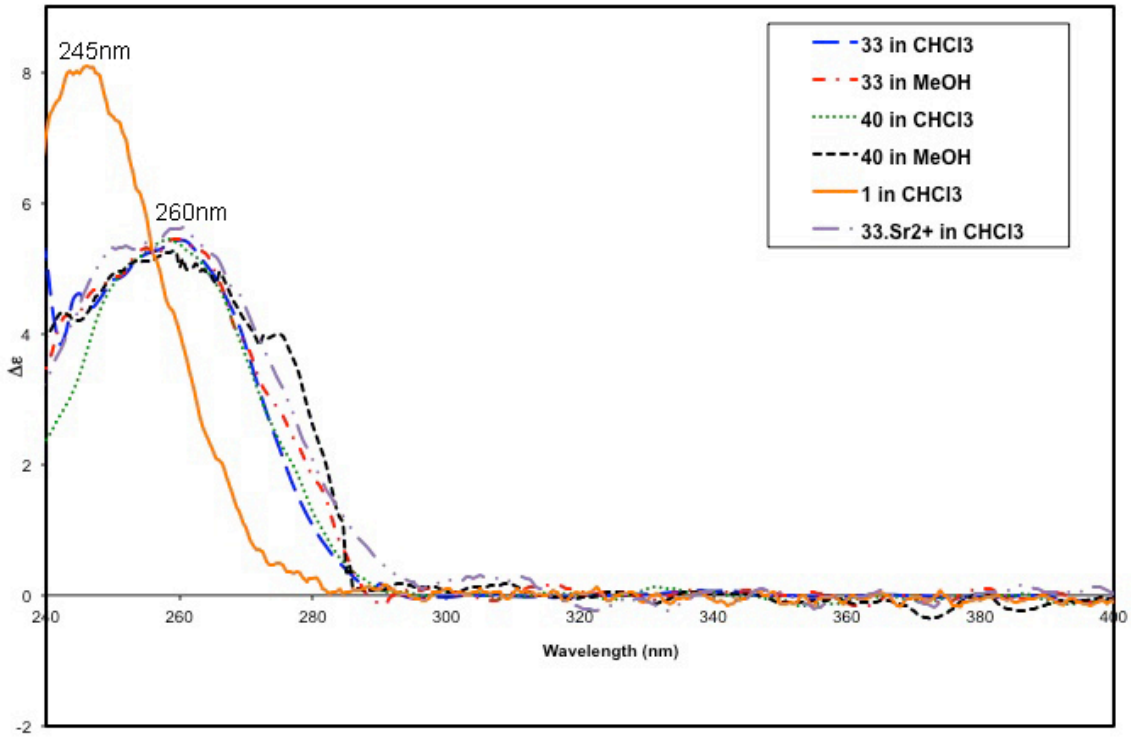


Figure 4.14 CD spectra of 33 and 34 in chloroform and methanol. (Spectrum of 1 in chloroform provided for reference)

4.5 Experimental section

4.5.1 General information

^{1}H NMR spectra were measured on a Bruker Avance 400 MHz in DMSO-d_6 [using DMSO (for 1H, δ = 2.50) as internal standard] or in $CDCl_3$ [using $CHCl_3$ (for ^{1}H, δ = 7.26) as internal standard]. ^{13}C NMR spectra were measured on a Bruker Avance 400 MHz spectrometer in DMSO-d_6 [using DMSO (for ^{13}C, δ = 39.5) as internal standard]. Chemical shifts are reported in ppm from tetramethylsilane. The following abbreviations are used to explain the multiplicities: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, td = triplet of doublets, q = quartet, m = multiplet, br = broad. 2D NOESY spectra were acquired with t_{mix} = 800 ms and d1 = 1500 ms. COSY spectra were acquired with d1 = 1500 ms. DOSY experiments were carried out on a Bruker Avance 400inv spectrometer equipped with a 5mm BBI Z-gradient probe (inverse broadband probe with z-gradient coil). All measurements were performed using the BPLED (ledbpgp2s) pulse sequence. The length of the diffusion gradient was optimized for each sample to obtain at least 95% signal attenuation due to diffusion. Δ and δ values respectively were found to be 60 ms and 3.6 ms (for **33** in $CDCl_3$), 60 ms and 4 ms (for **40** in $CDCl_3$), 60 ms and 3 ms (for **34** in $CDCl_3$), 60 ms and 4.4 ms (for **41** in $CDCl_3$), 65 ms and 10 ms (for **34**, **41** and **47** in DMSO-d_6) . Eddy current (t_e) was standardized at 5 ms. All measurements were taken at 298K with sample concentrations of 2.7 mM. Curve fits and diffusion coefficients were generated using the SimFit algorithm on the Bruker Topspin 2.1 software. MALDI-TOF mass analyses were performed on a Bruker Autoflex spectrometer in the reflectron mode using 2,5-dihydroxybenzoic acid (DHB) as the matrix. Flash column chromatography was performed using Silicycle 60 silica gel and eluting solvents were used

directly from their commercial bottles. Solvents and reagents for reactions were purchased commercially and used without further purification.

4.5.2 Synthesis of conjugates 33 and 40

To a stirred solution of cavitand **7** (42.1 mg, 0.0307 mmol) and nucleoside **35** (44.7 mg, 0.138 mmol) in argon-purged DMSO (3.8 mL) was added a solution of $CuSO_4 \bullet 5H_2O$ (102 μL, 0.04 M in argon-purged Milli-Q water, 0.00408 mmol) followed by a solution of sodium ascorbate (102 μL, 0.4 M in argon-purged Milli-Q water, 0.0408 mmol). The reaction was stirred at 60 °C for 20 h. The solvent was then removed *in vacuo* and the residue suspended in water. A few drops of ammonium hydroxide were added to remove the copper catalyst and the mixture was suction filtered. The residue, crude **33**, was washed with deionized water and allowed to suction dry, whereupon it was purified via step gradient flash chromatography ($CHCl_3$: MeOH 49 : 1 – $CHCl_3$: MeOH 19 : 1) to afford pure **33** in 44% yield (25.6 mg, 0.0096 mmol).

Compound 33: White glassy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 0.84 (12H, t, J = 6.4 Hz, CH_3 feet), 1.22 – 1.44 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 1.72 (12H, s, 5-Me), 2.31 (8H, br, -CH_2- feet), 4.26 (4H, d, J = 7.2 Hz, H_{in}), 4.38 (4H, m, H_4'), 4.55 (4H, t, J = 8.0 Hz, H_d), 4.65 (4H, dd, J = 6.8, 13.6 Hz, H_{5b}'), 4.74 (4H, dd, J = 4.8, 14.4 Hz, H_{5a}'), 4.87 – 4.89 (4H, m, H_3'), 4.93 (8H, s, H_a, H_b), 5.03 (4H, dd, J = 1.6, 6.4 Hz, H_2'), 5.74 (4H, s, H_1'), 5.90 (4H, d, J = 6.8 Hz, H_{out}), 7.27 (4H, s, ArH), 7.42 (4H, s, H6), 8.18 (4H, s, H_c), 11.43 (4H, s, NH); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.8, 150.3, 147.5, 143.8, 143.5, 138.7, 138.7, 124.9, 113.5, 110.0, 92.6, 84.7, 83.5, 81.3, 66.6, 51.1, 31.3, 29.2, 29.08, 29.06, 28.8, 27.7, 26.8, 25.1, 22.1, 13.9, 11.9; MS(MALDI-TOF): Found: *m/z* 2665.8. Calcd for $C_{140}H_{189}N_{20}O_{32}$: $(M+H)^+$ 2664.1.

The synthesis of **40** was carried out following the same procedure with the coupling of cavitand **7** to nucleoside **42**.

Compound 40: White powdery solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 0.85 (12H, t, J = 6.8 Hz, CH_3 feet), 1.24 – 1.46 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 1.78 (12H, s, 5-Me), 2.30 (8H, br, -CH_2- feet), 3.16 (12H, s, *N*-Me), 4.26 (4H, d, J = 5.2 Hz, H_{in}), 4.40 – 4.41 (4H, m, H_4'), 4.55 – 4.60 (4H, m, H_d), 4.65 – 4.79 (8H, m, H_{5a}', H_{5b}'), 4.93 – 4.95 (12H, m, H_3', H_a, H_b), 5.08 (4H, d, J = 5.2 Hz, H_2'), 5.80 (4H, s, H_1'), 5.89 (4H, d, J = 5.2 Hz, H_{out}), 7.27 (4H, s, ArH), 7.51 (4H, s, H6), 8.18 (4H, s, H_c); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.9, 150.4, 147.4, 143.8, 143.5, 138.7, 137.2, 125.0, 113.4, 108.6, 93.6, 85.0, 83.6, 81.3, 66.6, 51.1, 31.3, 29.2, 29.10, 29.06, 28.8, 27.7, 27.4, 26.8, 25.1, 22.1, 13.9, 12.5; MS(MALDI-TOF): Found: *m/z* 2720.6. Calcd for $C_{144}H_{197}N_{20}O_{32}$: $(M+H)^+$ 2720.2.

4.5.3 Synthesis of conjugates 34, 41 and 47

4.5.3.1 Preparation of 5'-OH dinucleoside 37

To a stirred solution of **36** (0.40 g, 0.56 mmol) in DMF (6 mL) was added TBAF (0.7 mL of a 1.0 M solution in THF, 0.67 mmol), and the mixture was stirred for 2 h. The solvent was then removed *in vacuo* and the residue was co-evaporated thrice with toluene. The crude material was purified by flash column chromatography ($CHCl_3$: EtOH 9 : 1) to afford **37** in 56% yield (0.19 g, 0.31 mmol).

Compound 37: White flaky solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.77 (3H, s, 5-Me), 1.78 (3H, s, 5-Me),

2.09 – 2.27 (2H, m, H_{2a}", H_{2b}"), 3.53 – 3.63 (2H, m, H_{5a}" and H_{5b}"), 3.95 – 3.97 (1H, m, H_4"), 4.20 – 4.22 (1H, m, H_3"), 4.35 – 4.39 (1H, m, H_4'), 4.57 (2H, s, H_a' and H_b'), 4.65 (1H, dd, J = 7.6, 14.0 Hz, H_{5b}'), 4.76 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.89 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.08 – 5.12 (2H, m, H_2', hydroxyl OH), 5.76 (1H, d, J = 1.6 Hz, H_1'), 6.11 (1H, dd, J = 5.6, 8.0 Hz, H_1"), 7.49 (1H, s, H6'), 7.70 (1H, s, H6"), 8.14 (1H, s, H_c'), 11.31 (1H, s, H2), 11.46 (1H, s, H1); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.9, 163.7, 150.5, 150.3, 143.9, 138.8, 136.0, 124.6, 113.5, 109.6, 109.5, 92.6, 84.9, 84.6, 83.8, 83.5, 81.3, 79.1, 61.7, 61.5, 51.2, 36.3, 26.9, 25.1, 12.3, 11.9; ESIHRMS: Found: *m/z* 604.2374. Calcd for $C_{26}H_{34}N_7O_{10}$: $(M+H)^+$ 604.2367.

4.5.3.2 Preparation of 5'-OH dinucleoside 44 (similar procedure used for 49)

To a stirred solution of **43** (1.13 g, 1.52 mmol) in THF (15 mL) was added TBAF (1.8 mL of a 1.0 M solution in THF, 1.82 mmol), and the mixture was stirred for 2 h. The solvent was then removed *in vacuo* and the residue re-dissolved in $CHCl_3$. The solvent was removed *in vacuo* once again, and the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 49 : 1 – $CHCl_3$: EtOH 19 : 1) to afford **44** in 96% yield (0.92 g, 1.46 mmol).

Compound 44: White flaky solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.83 (3H, s, 5-Me), 1.84 (3H, s, 5-Me), 2.12 – 2.33 (2H, m, H_{2a}", H_{2b}"), 3.17 (3H, s, N-Me), 3.18 (3H, s, N-Me), 3.54 – 3.64 (2H, m, H_{5a}" and H_{5b}"), 3.98 – 4.00 (1H, m, H_4"), 4.21 – 4.23 (1H, m, H_3"), 4.38 – 4.42 (1H, m, H_4'), 4.57 (2H, s, H_a' and H_b'), 4.66 (1H, dd, J = 7.6, 14.0 Hz, H_{5b}'), 4.77 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.92 (1H, dd, J = 4.0, 6.0 Hz, H_3'), 5.10 – 5.14 (2H, m, H_2', hydroxyl OH), 5.79 (1H, d, J = 1.6

Hz, H_1'), 6.16 (1H, dd, *J* = 6.0, 8.0 Hz, H_1"), 7.57 (1H, s, H6'), 7.78 (1H, s, H6"), 8.14 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 162.9, 162.8, 150.5, 150.4, 143.8, 137.3, 134.4, 124.7, 113.4, 108.5, 108.4, 93.6, 85.1, 84.9, 84.7, 83.6, 81.2, 78.8, 61.7, 61.4, 51.2, 36.6, 27.5, 27.4, 26.8, 25.1, 13.0, 12.5; ESIHRMS: Found: *m/z* 654.2493. Calcd for $C_{28}H_{37}N_7O_{10}Na$: $(M+Na)^+$ 654.2500.

Compound 49: White flaky solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.78 (3H, s, 5-Me), 1.82 (3H, s, 5-Me), 2.10 – 2.26 (2H, m, H_{2a}", H_{2b}"), 3.18 (3H, s, N-Me), 3.54 – 3.62 (2H, m, H_{5a}" and H_{5b}"), 3.95 (1H, d, *J* = 1.6 Hz, H_4"), 4.20 – 4.21 (1H, m, H_3"), 4.37 – 4.41 (1H, m, H_4'), 4.57 (2H, s, H_a' and H_b'), 4.66 (1H, dd, *J* = 7.6, 14.0 Hz, H_{5b}'), 4.78 (1H, dd, *J* = 4.8, 14.0 Hz, H_{5a}'), 4.92 (1H, dd, *J* = 4.4, 6.4 Hz, H_3'), 5.11 – 5.12 (2H, m, H_2', hydroxyl OH), 5.79 (1H, d, *J* = 1.6 Hz, H_1'), 6.11 (1H, dd, *J* = 6.0, 8.4 Hz, H_1"), 7.57 (1H, s, H6'), 7.70 (1H, s, H6"), 8.14 (1H, s, H_c'), 11.30 (1H, s, H2); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.7, 163.0, 150.4, 143.8, 137.4, 135.9, 124.7, 113.5, 109.5, 108.6, 93.7, 85.2, 84.6, 83.8, 83.6, 81.3, 79.1, 61.8, 61.5, 51.2, 36.3, 27.4, 26.9, 25.1, 12.6, 12.3; ESIHRMS: Found: *m/z* 640.2341. Calcd for $C_{27}H_{35}N_7O_{10}Na$: $(M+Na)^+$ 640.2343.

4.5.3.3 Preparation of 5'-I dinucleoside 38

To a stirred solution of **37** (0.19 g, 0.31 mmol) in DMF (3 mL) was added PPh_3 (0.27 g, 1.04 mmol), imidazole (0.14 g, 2.08 mmol) and I_2 (0.25 g, 1.0 mmol, 4 portions over 5 mins) in that order. The reaction was stirred at room temperature overnight, whereupon the solvent was removed *in vacuo*. H_2O (10 mL) was added to the residue and the mixture was extracted with $CHCl_3$ (3 x 50 mL). The combined organic extracts were washed with brine and dried over

anhydrous $MgSO_4$. After evaporation of the solvent, the crude material was reprecipitated from $CHCl_3$ to afford **38** in 89% yield (0.2 g, 0.28 mmol).

Compound 38: White solid; ^1H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.77 (3H, s, 5-Me), 1.81 (3H, s, 5-Me), 2.23 – 2.39 (2H, m, H_{2a}" and H_{2b}"), 3.40 – 3.54 (2H, m, H_{5a}", H_{5b}"), 4.04 – 4.07 (1H, m, H_4"), 4.16 – 4.18 (1H, m, H_3"), 4.36 – 4.40 (1H, m, H_4'), 4.62 – 4.68 (3H, m, H_a', H_b', H_{5b}'), 4.77 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.89 (1H, dd, J = 4.0, 6.0 Hz, H_3'), 5.09 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.76 (1H, d, J = 2.0 Hz, H_1'), 6.16 (1H, dd, J = 6.0, 8.8 Hz, H_1"), 7.50 (1H, s, H6'), 7.56 (1H, s, H6"), 8.17 (1H, s, H_c'), 11.36 (1H, s, H2), 11.48 (1H, s, H1); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.9, 163.6, 150.4, 150.3, 143.7, 138.9, 136.0, 124.7, 113.5, 109.9, 109.6, 92.6, 84.9, 84.3, 83.5, 82.9, 81.3, 80.9, 61.9, 51.2, 35.0, 26.9, 25.1, 12.1, 12.0, 7.6; ESIHRMS: Found: *m/z* 714.1379. Calcd for $C_{26}H_{33}N_7O_9I$: $(M+H)^+$ 714.1385.

4.5.3.4 Preparation of 5'-I dinucleoside 45 (similar procedure used for 50)

To a stirred solution of **44** (0.92 g, 1.46 mmol) in DMF (6 mL) was added PPh_3 (1.26 g, 4.81 mmol), imidazole (0.66 g, 9.62 mmol) and I_2 (1.18 g, 4.67 mmol, 4 portions over 5 mins) in that order. The reaction was stirred at room temperature overnight, whereupon the solvent was removed *in vacuo*. H_2O (10 mL) was added to the residue and the mixture was extracted with $CHCl_3$ (3 x 50 mL). The combined organic extracts were washed with brine and dried over anhydrous $MgSO_4$. After evaporation of the solvent, the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 49 : 1 – $CHCl_3$: EtOH 24 : 1) to provide **45** in 60% yield (0.64 g, 0.86 mmol).

Compound 45: White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.83 (3H, s, 5-Me), 1.86 (3H, s, 5-Me), 2.25 – 2.40 (2H, m, H_{2a}" and H_{2b}"), 3.17 (3H, s, N-Me), 3.19 (3H, s, N-Me), 3.40 – 3.55 (2H, m, H_{5a}", H_{5b}"), 4.08 (1H, td, J = 2.0, 6.4, 6.4 Hz, H_4"), 4.17 – 4.18 (1H, m, H_3"), 4.38 – 4.42 (1H, m, H_4'), 4.59 – 4.69 (3H, m, H_a', H_b', H_{5b}'), 4.78 (1H, dd, J = 4.8, 14.4 Hz, H_{5a}'), 4.92 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.11 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.79 (1H, d, J = 1.6 Hz, H_1'), 6.21 (1H, dd, J = 6.0, 8.4 Hz, H_1"), 7.58 (1H, s, H6'), 7.61 (1H, s, H6"), 8.17 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.0, 162.7, 150.6, 150.4, 143.6, 137.4, 134.5, 124.9, 113.5, 108.9, 108.6, 93.7, 85.4, 85.2, 83.6, 83.1, 81.3, 80.8, 61.9, 51.2, 35.2, 27.6, 27.5, 26.9, 25.1, 12.8, 12.6, 7.5; ESIHRMS: Found: *m/z* 742.1694. Calcd for $C_{28}H_{37}N_7O_9I$: $(M+H)^+$ 742.1698.

Compound 50: Yield: 86%. White flakes; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.81 (3H, s, 5-Me), 1.83 (3H, s, 5-Me), 2.22 – 2.39 (2H, m, H_{2a}" and H_{2b}"), 3.18 (3H, s, N-Me), 3.39 – 3.54 (2H, m, H_{5a}", H_{5b}"), 4.05 (1H, td, J = 1.6, 6.4, 6.4 Hz, H_4"), 4.16 – 4.17 (1H, m, H_3"), 4.37 – 4.42 (1H, m, H_4'), 4.61 – 4.69 (3H, m, H_a', H_b', H_{5b}'), 4.78 (1H, dd, J = 4.8, 14.4 Hz, H_{5a}'), 4.92 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.11 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.79 (1H, d, J = 1.6 Hz, H_1'), 6.16 (1H, dd, J = 6.0, 8.4 Hz, H_1"), 7.56 (1H, s, H6'), 7.58 (1H, s, H6"), 8.17 (1H, s, H_c'), 11.37 (1H, s, H2); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.6, 162.9, 150.4, 143.6, 137.4, 136.0, 124.8, 113.5, 109.9, 108.6, 93.7, 85.2, 84.3, 83.6, 82.9, 81.3, 80.9, 61.9, 51.2, 35.0, 27.5, 26.9, 25.1, 12.6, 12.1, 7.6; ESIHRMS: Found: *m/z* 750.1368. Calcd for $C_{27}H_{34}N_7O_9NaI$: $(M+Na)^+$ 750.1360.

4.5.3.5 Preparation of 5'-N_3 dinucleoside 39

To a stirred solution of **38** (0.2 g, 0.28 mmol) in DMF (2.5 mL) was added NaN_3 (37 mg, 0.56 mmol). The reaction was then stirred at 80 ºC for 3 h, whereupon the solvent was removed *in vacuo*. H_2O was added and the mixture was extracted with EtOAc (3 x 25 mL). The combined organic extracts were washed with brine and dried over anhydrous $MgSO_4$. After evaporation of the solvent, the residue was taken up in cold acetone, triturated, suction filtered and washed twice with cold acetone. The resultant white powder **39** was collected (61.5 mg, 0.1 mmol, 36% yield) and pure enough (vide NMR) for use.

Compound 39: White solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.30 (3H, s, isopropylidene CH_3), 1.49 (3H, s, isopropylidene CH_3), 1.78 (3H, s, 5-Me), 1.81 (3H, s, 5-Me), 2.24 – 2.37 (2H, m, H_{2a}" and H_{2b}"), 3.56 – 3.64 (2H, m, H_{5a}" and H_{5b}"), 4.06 – 4.07 (1H, m, H_4"), 4.18 – 4.19 (1H, m, H_3"), 4.38 – 4.39 (1H, m, H_4'), 4.57 – 4.69 (3H, m, H_{5b}', H_a', H_b'), 4.77 (1H, dd, J = 5.2, 14.0 Hz, H_{5a}'), 4.90 (1H, t, J = 4.4 Hz, H_3'), 5.10 (1H, d, J = 6.4 Hz, H_2'), 5.77 (1H, s, H_1'), 6.16 (1H, t, J = 6.4 Hz, H_1"), 7.51 (1H, s, H6'), 7.54 (1H, s, H6"), 8.16 (1H, s, H_c'), 11.37 (1H, s, H2), 11.48 (1H, s, H1); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.9, 163.6, 150.5, 150.3, 143.7, 138.9, 136.0, 124.7, 113.5, 109.9, 109.6, 92.6, 84.9, 84.2, 83.5, 82.1, 81.3, 78.9, 62.0, 51.9, 51.2, 35.2, 26.9, 25.1, 12.1, 11.9; ESIHRMS: Found: *m/z* 651.2247. Calcd for $C_{26}H_{32}N_{10}O_9Na$: $(M+Na)^+$ 651.2251.

4.5.3.6 Preparation of 5'-N_3 dinucleoside 46 (similar procedure used for 51)

To a stirred solution of **45** (0.63 g, 0.84 mmol) in DMF (3 mL) was added NaN_3 (0.11 g, 1.69 mmol). The reaction was then stirred at 80 ºC for 3 h, whereupon the solvent was removed

in vacuo. H_2O was added and the mixture was extracted with $CHCl_3$ (3 x 50 mL). The combined organic extracts were washed with H_2O, brine and dried over anhydrous $MgSO_4$. After evaporation of the solvent, the crude material was purified by flash column chromatography using step gradient elution ($CHCl_3$: EtOH 49 : 1 – $CHCl_3$: EtOH 47 : 3) to provide **46** in 84% yield (0.47 g, 0.71 mmol).

Compound 46: White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.83 (3H, s, 5-Me), 1.86 (3H, s, 5-Me), 2.25 – 2.37 (2H, m, H_{2a}" and H_{2b}"), 3.17 (3H, s, N-Me), 3.18 (3H, s, N-Me), 3.54 – 3.64 (2H, m, H_{5a}" and H_{5b}"), 4.05 – 4.09 (1H, m, H_4"), 4.17 – 4.19 (1H, m, H_3"), 4.37 – 4.41 (1H, m, H_4'), 4.59 (2H, s, H_a' and H_b'), 4.66 (1H, dd, J = 7.6, 14.0 Hz, H_{5b}'), 4.78 (1H, dd, J = 4.4, 14.0 Hz, H_{5a}'), 4.92 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.11 (1H, dd, J = 1.6, 6.4 Hz, H_2'), 5.79 (1H, d, J = 2.0 Hz, H_1'), 6.19 (1H, t, J = 6.4 Hz, H_1"), 7.58 (2H, s, H6', H6"), 8.15 (1H, s, H_c'); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.0, 162.7, 150.6, 150.4, 143.6, 137.4, 134.5, 124.8, 113.5, 108.8, 108.6, 93.7, 85.3, 85.2, 83.6, 82.2, 81.3, 78.9, 62.1, 51.8, 51.2, 35.4, 27.6, 27.4, 26.8, 25.1, 12.8, 12.6; ESIHRMS: Found: *m/z* 679.2567. Calcd for $C_{28}H_{36}N_{10}O_9Na$: $(M+Na)^+$ 679.2564.

Compound 51: Yield: 95%. White foamy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 1.29 (3H, s, isopropylidene CH_3), 1.48 (3H, s, isopropylidene CH_3), 1.80 (3H, s, 5-Me), 1.82 (3H, s, 5-Me), 2.22 – 2.35 (2H, m, H_{2a}" and H_{2b}"), 3.18 (3H, s, N-Me), 3.53 – 3.62 (2H, m, H_{5a}" and H_{5b}"), 4.03 – 4.07 (1H, m, H_4"), 4.16 – 4.17 (1H, m, H_3"), 4.37 – 4.41 (1H, m, H_4'), 4.59 (2H, s, H_a' and H_b'), 4.66 (1H, dd, J = 8.0, 14.4 Hz, H_{5b}'), 4.78 (1H, dd, J = 4.8, 14.0 Hz, H_{5a}'), 4.92 (1H, dd, J = 4.4, 6.4 Hz, H_3'), 5.11 (1H, dd, J = 2.0, 6.8 Hz, H_2'), 5.79 (1H, d, J = 1.6 Hz, H_1'), 6.14 (1H, t, J = 6.0 Hz, H_1"), 7.53 (1H, s, H6'), 7.58 (1H, s, H6"), 8.16 (1H, s, H_c'), 11.37 (1H, s, H2); ^{13}C (100 MHz,

DMSO-d_6) δ(ppm) 163.6, 162.9, 150.4, 143.6, 137.4, 136.0, 124.8, 113.5, 109.9, 108.6, 93.7, 85.2, 84.1, 83.6, 82.1, 81.3, 78.9, 62.0, 51.9, 51.2, 35.2, 27.4, 26.8, 25.1, 12.6, 12.1; ESIHRMS: Found: *m/z* 665.2416. Calcd for $C_{27}H_{34}N_{10}O_9Na$: $(M+Na)^+$ 665.2408.

4.5.3.7 Preparation of conjugate 34 (similar procedure used for 41 and 47)

To a stirred solution of cavitand **7** (25.7 mg, 0.0188 mmol) and nucleoside **39** (53.4 mg, 0.085 mmol) in argon-purged DMSO (2.3 mL) was added a solution of $CuSO_4•5H_2O$ (63 μL, 0.04 M in argon-purged Milli-Q water, 0.0025 mmol) followed by a solution of cesium ascorbate, freshly prepared from the stoichiometric reaction between cesium carbonate and ascorbic acid (63 μL, 0.4 M in argon-purged Milli-Q water, 0.025 mmol). The reaction was stirred at 80 °C for 24 h. The solvent was then removed *in vacuo* and the residue suspended in water. A few drops of ammonium hydroxide were added to remove the copper catalyst and the mixture was suction filtered. The residue, crude **34**, was washed with deionized water and allowed to suction dry, whereupon it was purified with flash column chromatography ($CHCl_3$: MeOH 9 : 1) to afford pure **34** in 37% yield (26.8 mg, 0.007 mmol).

Compound 34: White glassy solid; ^{1}H (400 MHz, DMSO-d_6) δ(ppm) 0.84 (12H, t, J = 6.4 Hz, CH_3 feet), 1.23 – 1.43 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 1.74 (12H, s, 5-Me'), 1.77 (12H, s, 5-Me''), 2.16 – 2.29 (16H, m, -CH_2 feet, H_{2a}'', H_{2b}''), 4.28 – 4.39 (16H, m, H_{in}, H_3'', H_4', H_4''), 4.53 – 4.77 (28H, m, H_d, H_a', H_b', H_{5a}', H_{5b}', H_{5a}'', H_{5b}''), 4.88 (4H, dd, J = 4.0, 6.0 Hz, H_3'), 4.97 (8H, s, H_a and H_b), 5.07 (4H, dd, J = 1.6, 6.4 Hz, H_2'), 5.74 (4H, d, J = 1.6 Hz, H_1'), 5.90 (4H, d, J = 5.6 Hz, H_{out}), 6.10 (4H, t, J = 7.2 Hz, H_1''), 7.27 (4H, s, ArH), 7.36 (4H, s, H6''), 7.47 (4H, s, H6'), 8.11 (4H, s, H_c'), 8.17 (4H, s, H_c''), 11.31 (4H, s, H_{inner}), 11.45 (4H, s,

H_{outer}); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.8, 163.5, 150.4, 150.3, 147.5, 143.7, 143.6, 143.5, 138.8, 138.7, 135.8, 125.0, 124.6, 113.4, 109.9, 109.5, 92.6, 84.8, 84.3, 83.5, 81.5, 81.3, 79.0, 62.0, 51.2, 31.4, 29.3, 29.20, 29.15, 29.0, 28.7, 27.7, 26.7, 25.0, 22.1, 13.9, 12.0, 11.9; MS(MALDI-TOF): Found: *m/z* 3887.5. Calcd for $C_{192}H_{249}N_{40}O_{48}$: $(M+H)^+$ 3885.3.

Compound 41: Yield: 42%. White glassy solid; ^1H (400 MHz, DMSO-d_6) δ(ppm) 0.84 (12H, t, J = 6.4 Hz, CH_3 feet), 1.23 – 1.43 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 1.79 (12H, s, 5-Me'), 1.80 (12H, s, 5-Me''), 2.19 – 2.35 (16H, m, CH_2 feet, H_{2a}'', H_{2b}''), 3.12 (12H, s, N-Me), 3.16 (12H, s, N-Me), 4.25 – 4.41 (16H, m, H_{in}, H_3'', H_4', H_4''), 4.51 – 4.79 (28H, m, H_d, H_a', H_b', H_{5a}', H_{5b}', H_{5a}'', H_{5b}''), 4.90 – 4.97 (12H, m, H_3', H_a, H_b), 5.09 (4H, d, J = 6.4 Hz, H_2'), 5.77 (4H, s, H_1'), 5.88 (4H, d, J = 6.0 Hz, H_{out}), 6.14 (4H, t, J = 6.4 Hz, H_1''), 7.26 (4H, s, ArH), 7.41 (4H, s, H6''), 7.54 (4H, s, H6'), 8.10 (4H, s, H_c'), 8.15 (4H, s, H_c''); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.4, 163.2, 150.94, 150.89, 147.9, 144.2, 144.0, 139.2, 137.8, 134.9, 125.5, 125.2, 113.9, 109.3, 109.0, 94.3, 86.1, 86.02, 85.98, 85.7, 84.1, 82.2, 81.94, 81.86, 81.8, 79.4, 62.6, 51.7, 40.9, 31.8, 29.8, 29.7, 29.58, 29.55, 29.3, 28.2, 27.92, 27.88, 27.2, 25.5, 22.6, 14.4, 13.1; MS(MALDI-TOF): Found: *m/z* 4000.5. Calcd for $C_{200}H_{265}N_{40}O_{48}$: $(M+H)^+$ 3997.5.

Compound 47: Yield: 83%. White solid; ^1H (400 MHz, DMSO-d_6) δ(ppm) 0.85 (12H, t, J = 6.4 Hz, CH_3 feet), 1.23 – 1.43 (96H, m, long chain aliphatic feet and isopropylidene CH_3), 1.77 (12H, s, 5-Me''), 1.80 (12H, s, 5-Me'), 2.16 – 2.28 (16H, m, CH_2 feet, H_{2a}'', H_{2b}''), 3.16 (12H, s, N-Me), 4.27 – 4.41 (16H, m, H_{in}, H_3'', H_4', H_4''), 4.55 – 4.79 (28H, m, H_d, H_a', H_b', H_{5a}', H_{5b}', H_{5a}'', H_{5b}''), 4.91 (4H, dd, J = 4.4, 6.0 Hz, H_3'), 4.98 (8H, s, H_a, H_b), 5.09 (4H, dd, J = 1.2, 6.4 Hz, H_2'), 5.78 (4H, s, H_1'), 5.90 (4H, d, J = 5.2 Hz, H_{out}), 6.10 (4H, t, J = 7.2 Hz, H_1''), 7.27 (4H, s, ArH), 7.36 (4H, s, H6''), 7.54 (4H, s, H6'), 8.09 (4H, s, H_c'), 8.16 (4H, s, H_c''), 11.31 (4H, s,

H_{inner}); ^{13}C (100 MHz, DMSO-d_6) δ(ppm) 163.6, 162.9, 150.4, 150.3, 147.5, 143.7, 143.5, 138.7, 137.4, 135.9, 125.1, 124.7, 113.4, 109.9, 108.6, 93.8, 85.2, 84.4, 83.6, 81.6, 81.3, 79.0, 62.0, 51.2, 31.3, 29.31, 29.26, 29.1, 28.8, 27.7, 27.4, 26.8, 25.0, 22.1, 13.9, 12.6, 12.0; MS(MALDI-TOF): Found: *m/z* 3940.9. Calcd for $C_{196}H_{257}N_{40}O_{48}$: $(M+H)^+$ 3941.4.

4.5.4 Supplementary ^1H and ^1H-^1H COSY spectra

Spectra begin on the following page.

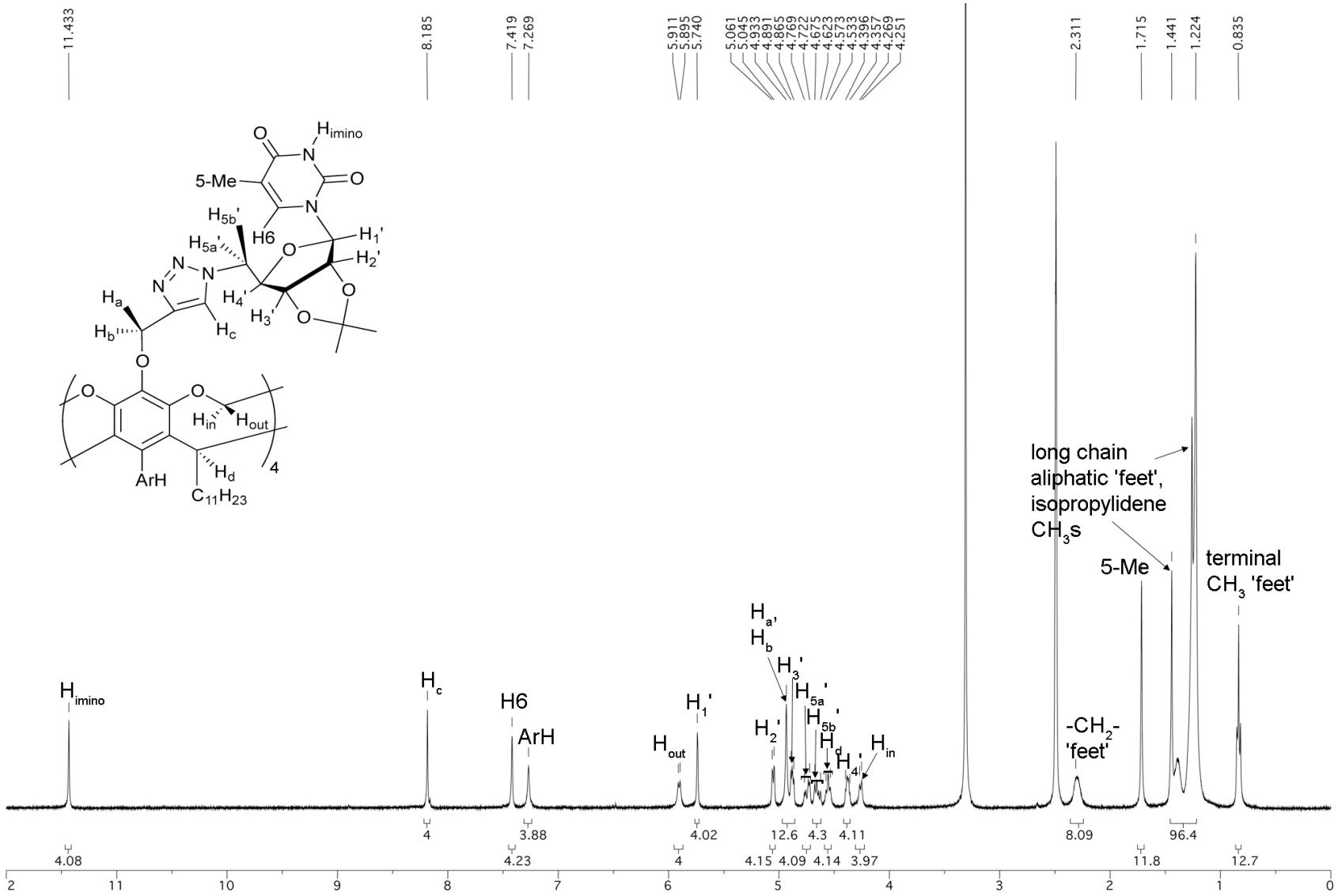


Figure 4.15 ^{1}H NMR spectrum of 33 in DMSO-d_6 at 25 ºC.

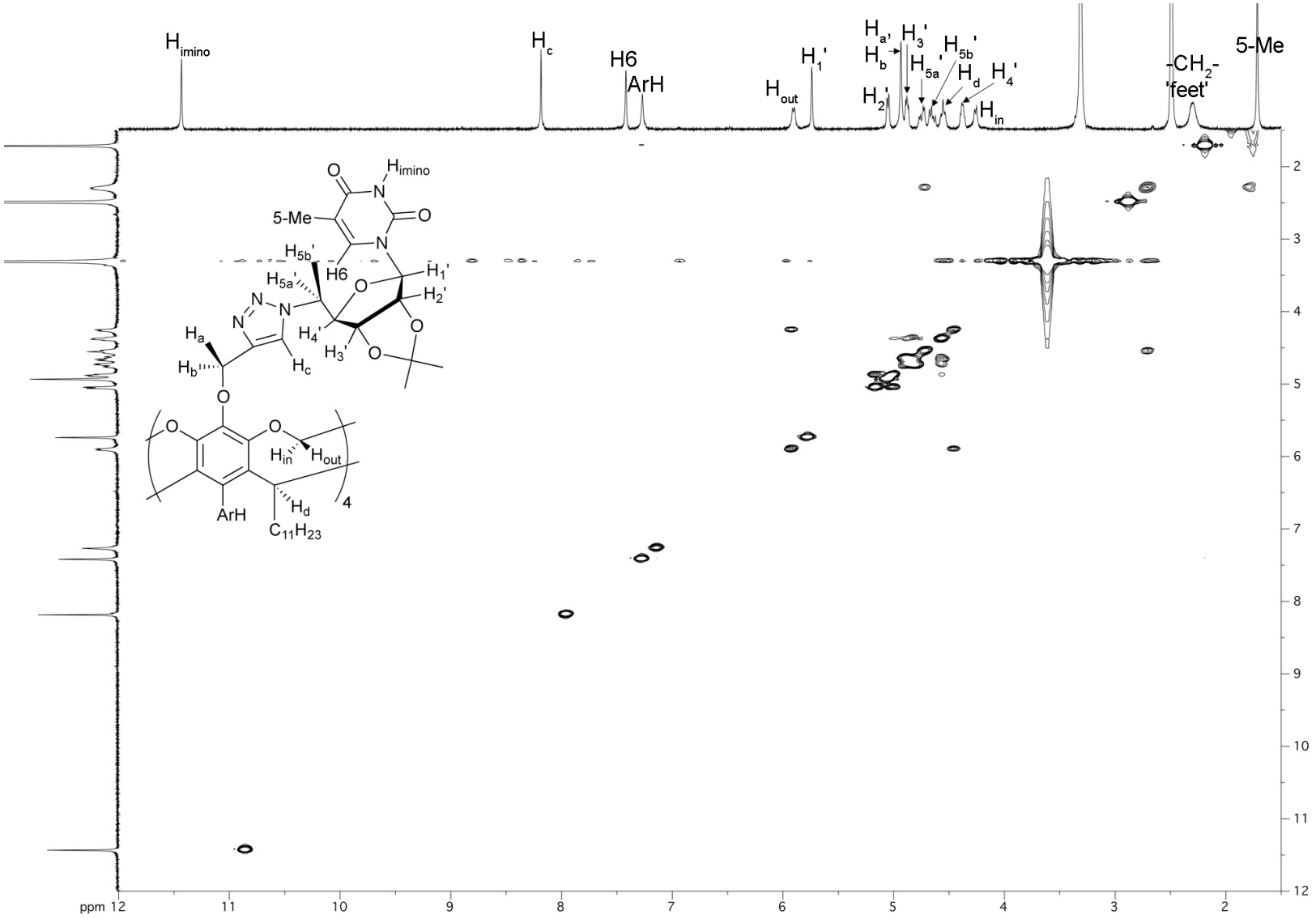


Figure 4.16 ^{1}H-^{1}H COSY spectrum of 33 in DMSO-d_6 at 25 ºC.

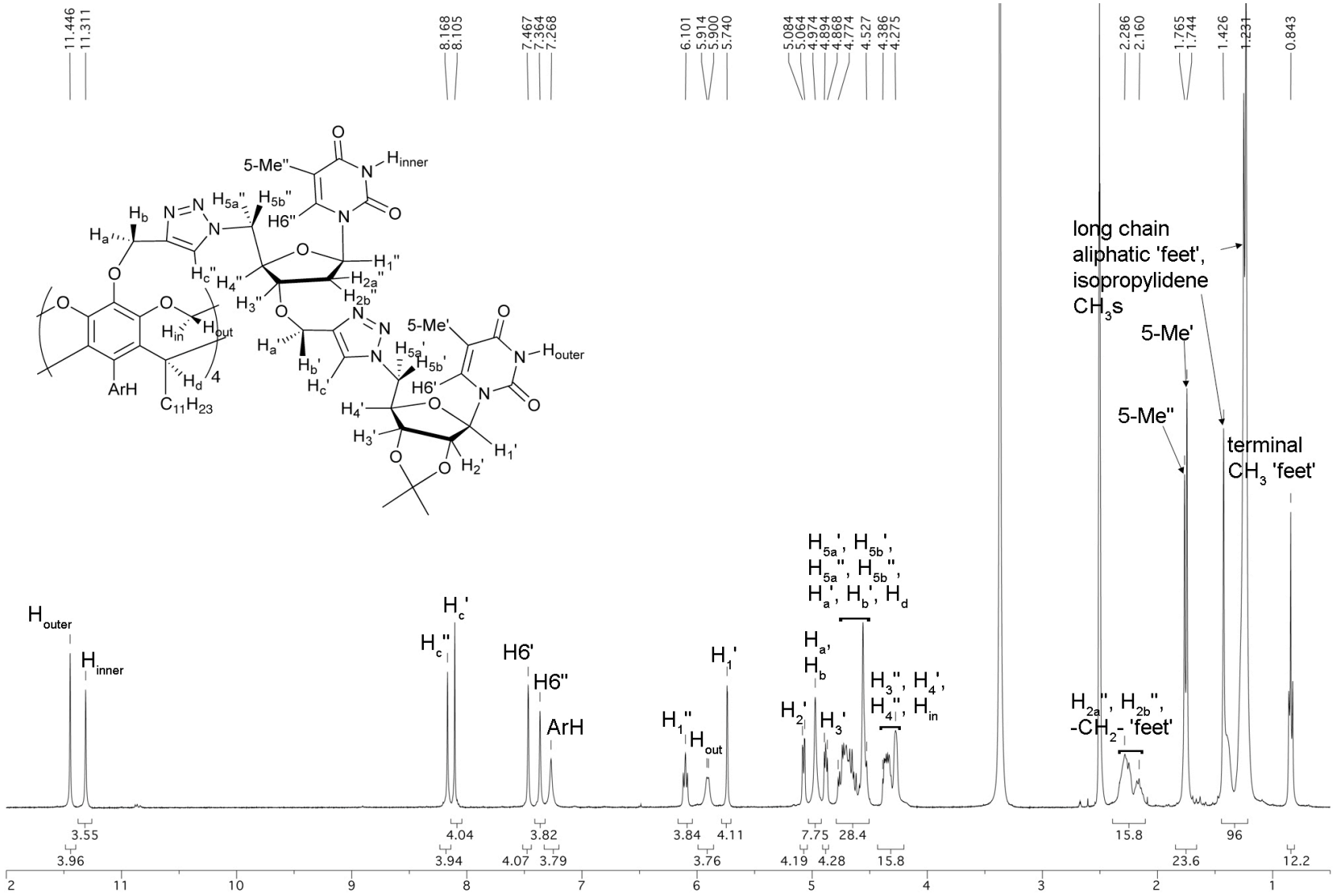


Figure 4.17 ^{1}H NMR spectrum of 34 in DMSO-d_6 at 25 ºC.

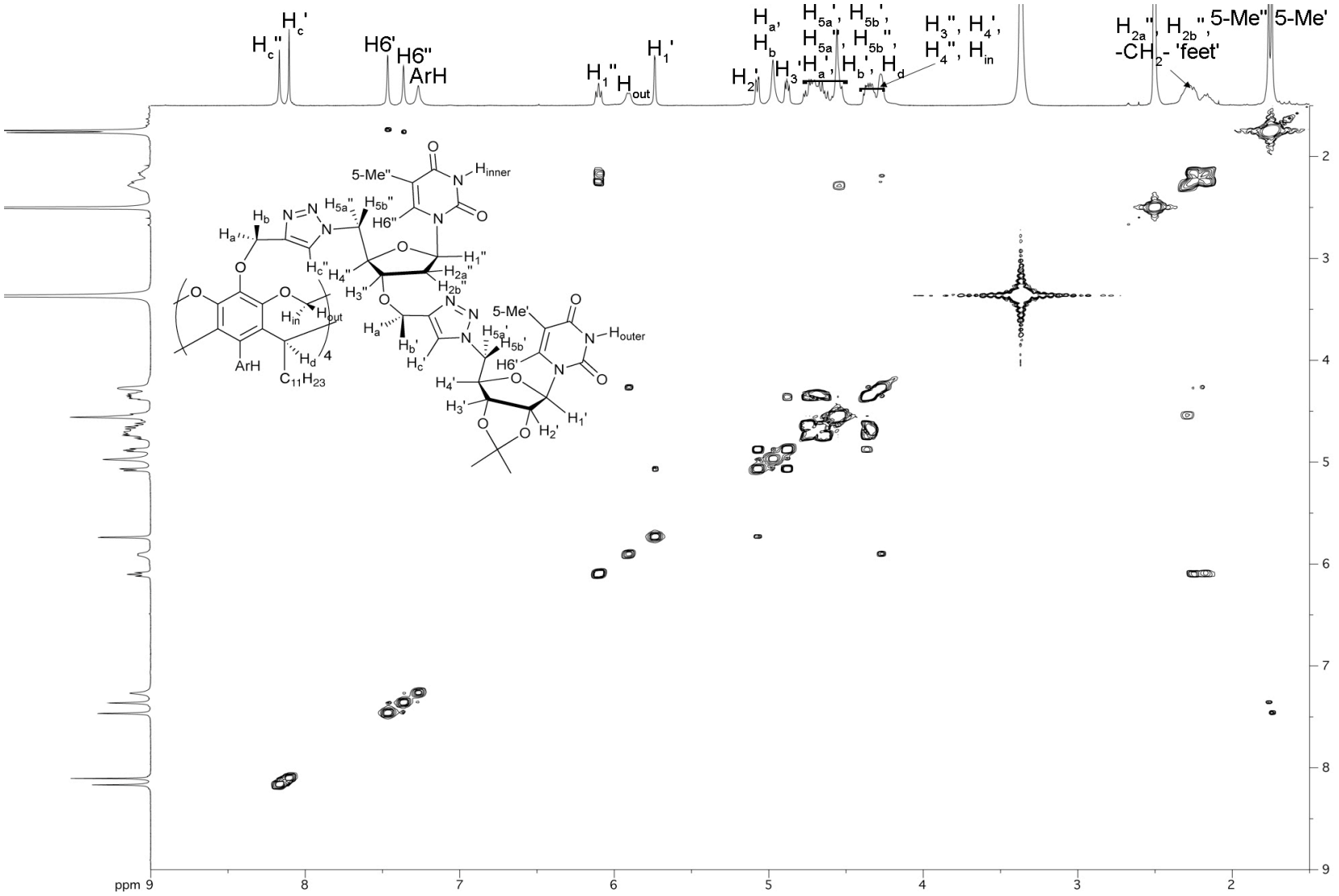


Figure 4.18 ^{1}H-^{1}H COSY spectrum of 34 in DMSO-d_6 at 25 ºC.

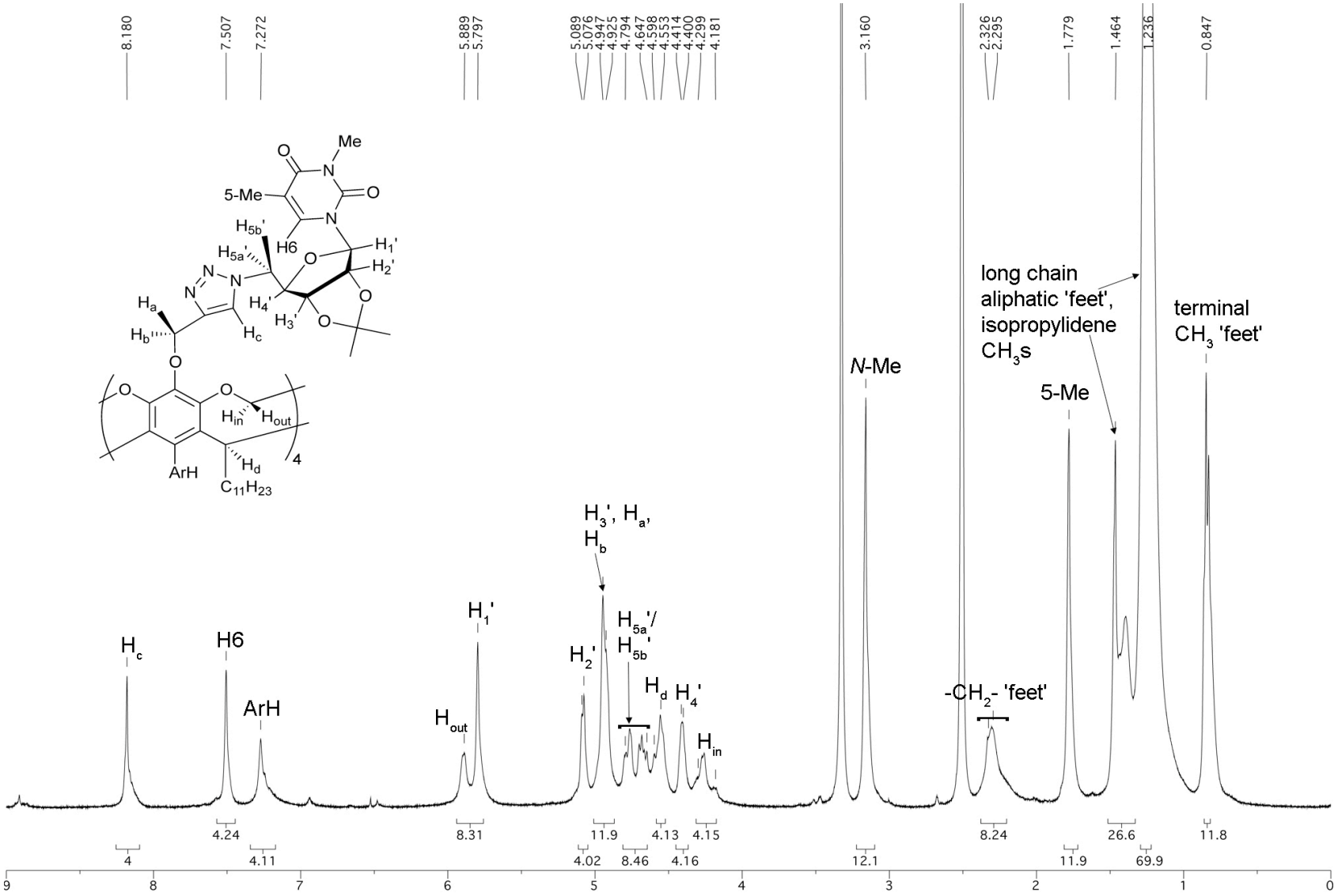


Figure 4.19 ^{1}H NMR spectrum of 40 in DMSO-d_6 at 25 °C.

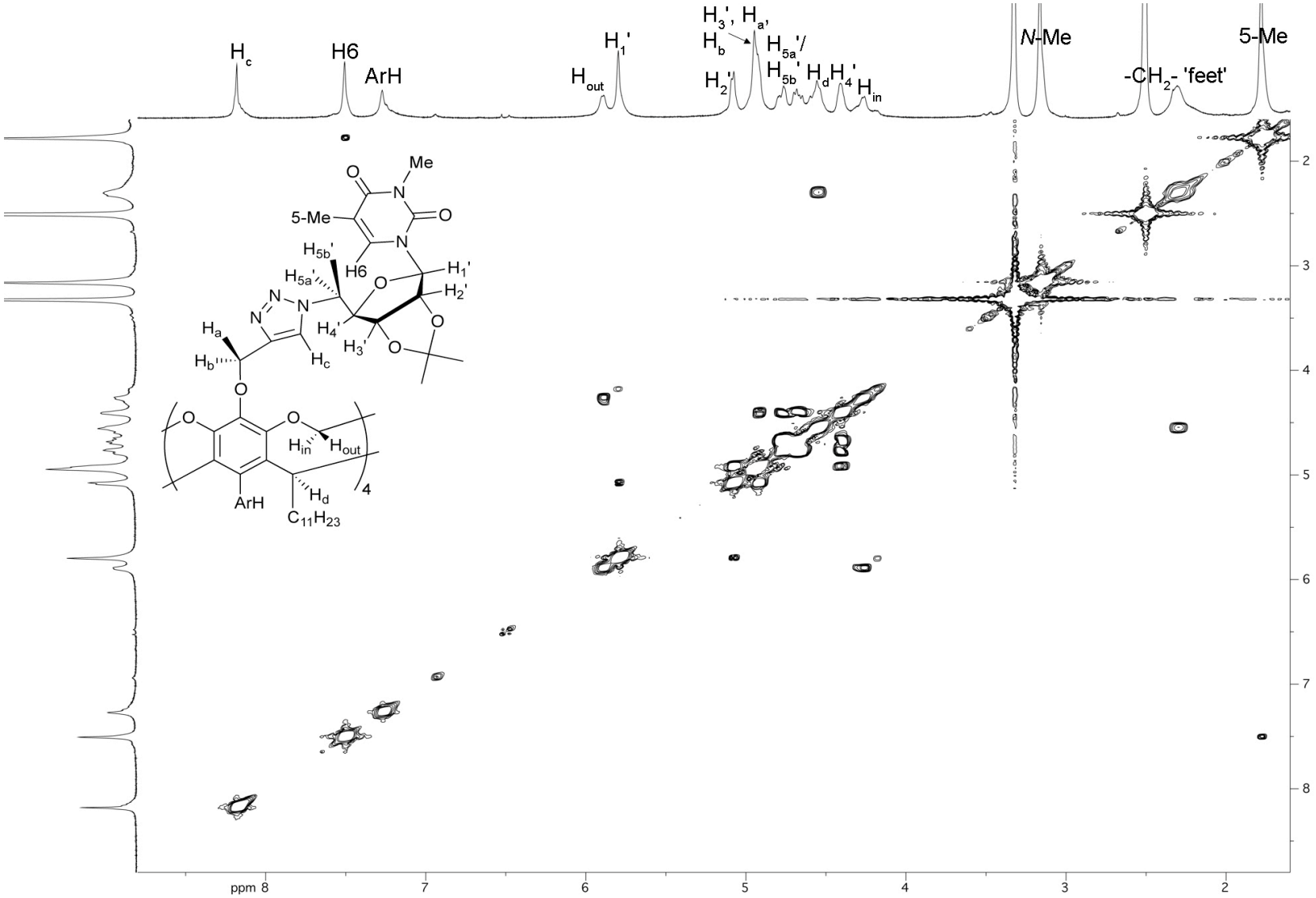


Figure 4.20 ^{1}H-^{1}H COSY spectrum of 40 in DMSO-d_6 at 25 ºC.

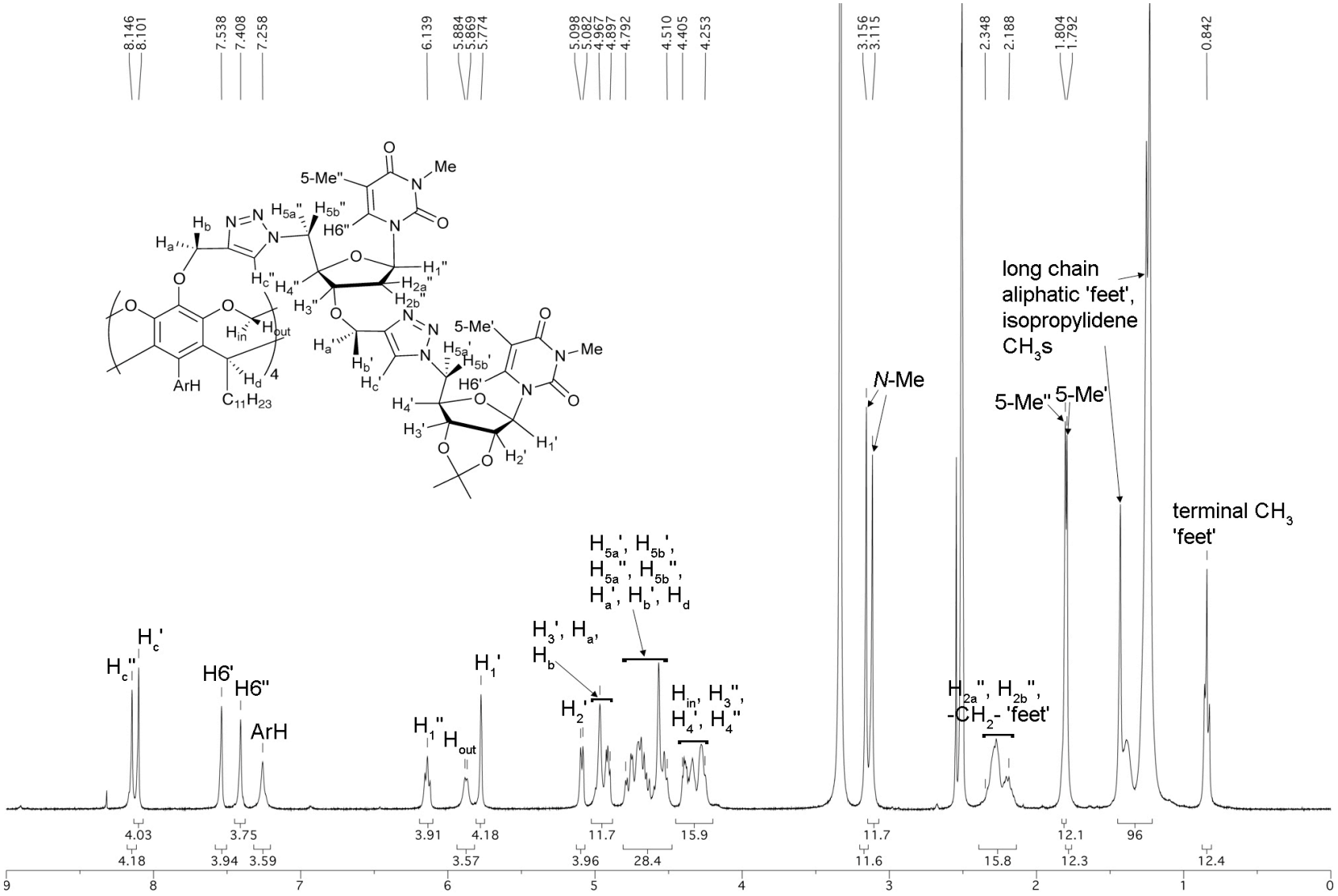


Figure 4.21 ^{1}H NMR spectrum of 41 in DMSO-d_6 at 25 ºC.

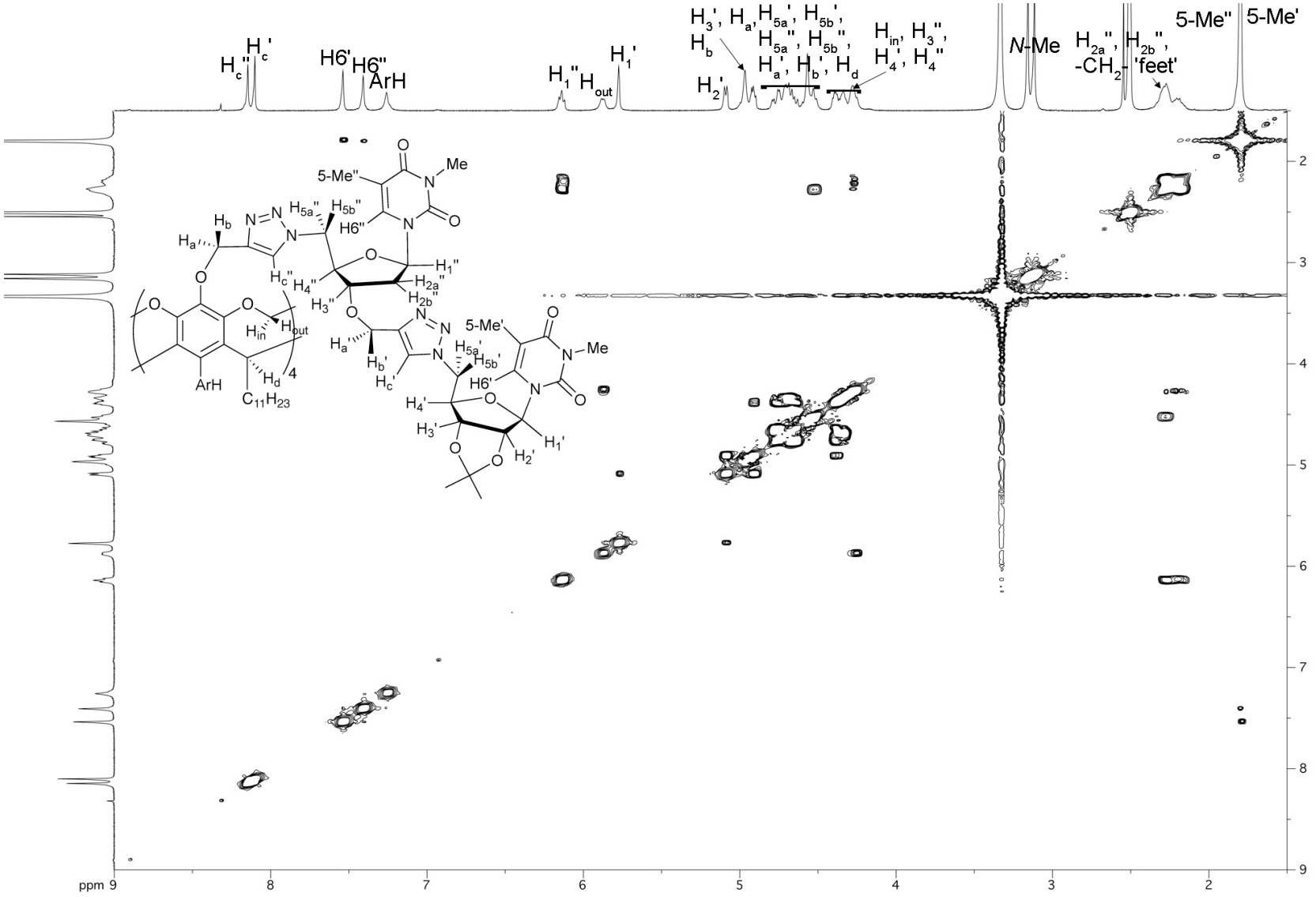


Figure 4.22 ^{1}H-^{1}H COSY spectrum of 41 in DMSO-d_6 at 25 ºC.

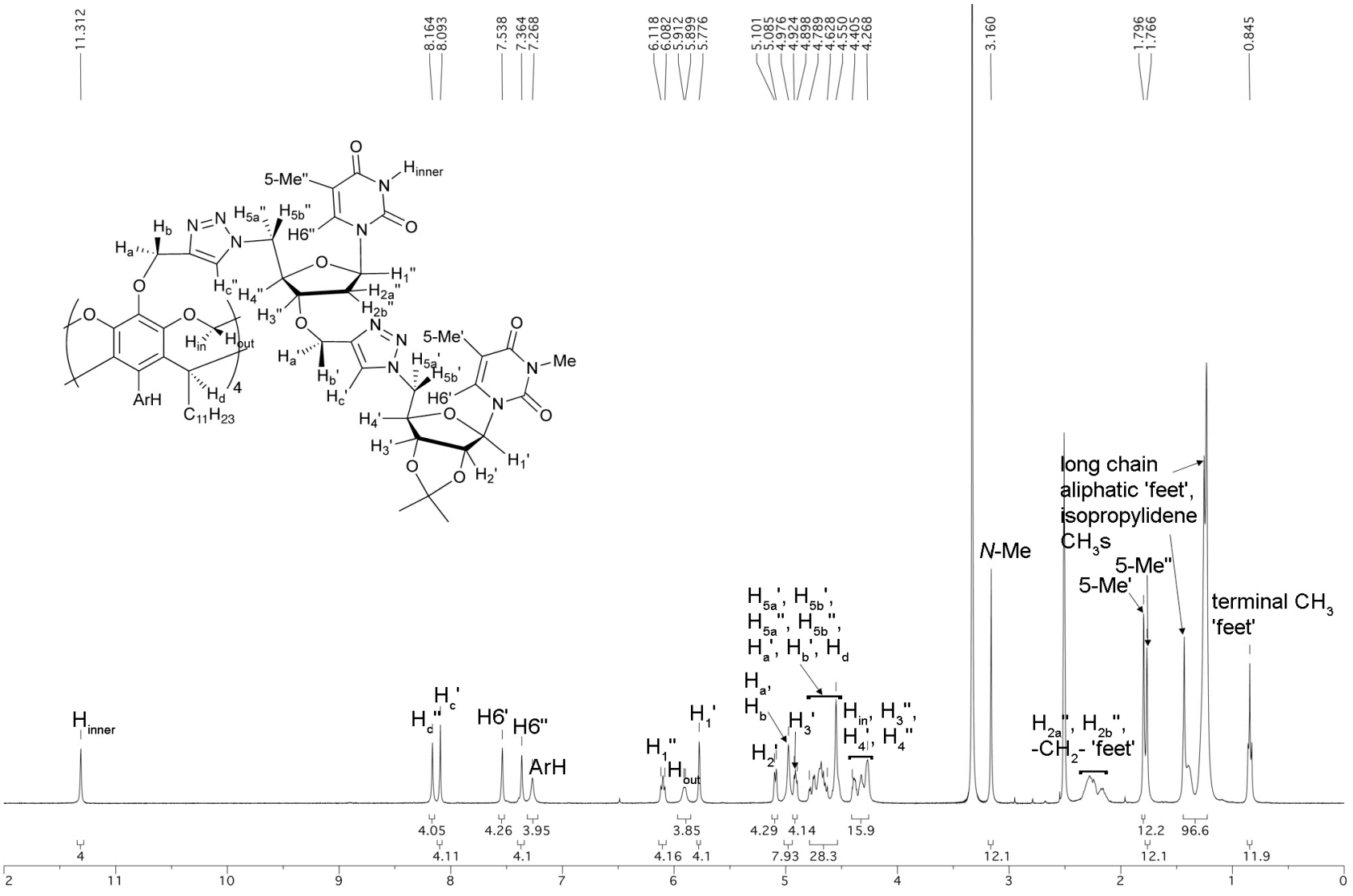


Figure 4.23 ^{1}H NMR spectrum of 47 in DMSO-d_6 at 25 ºC.

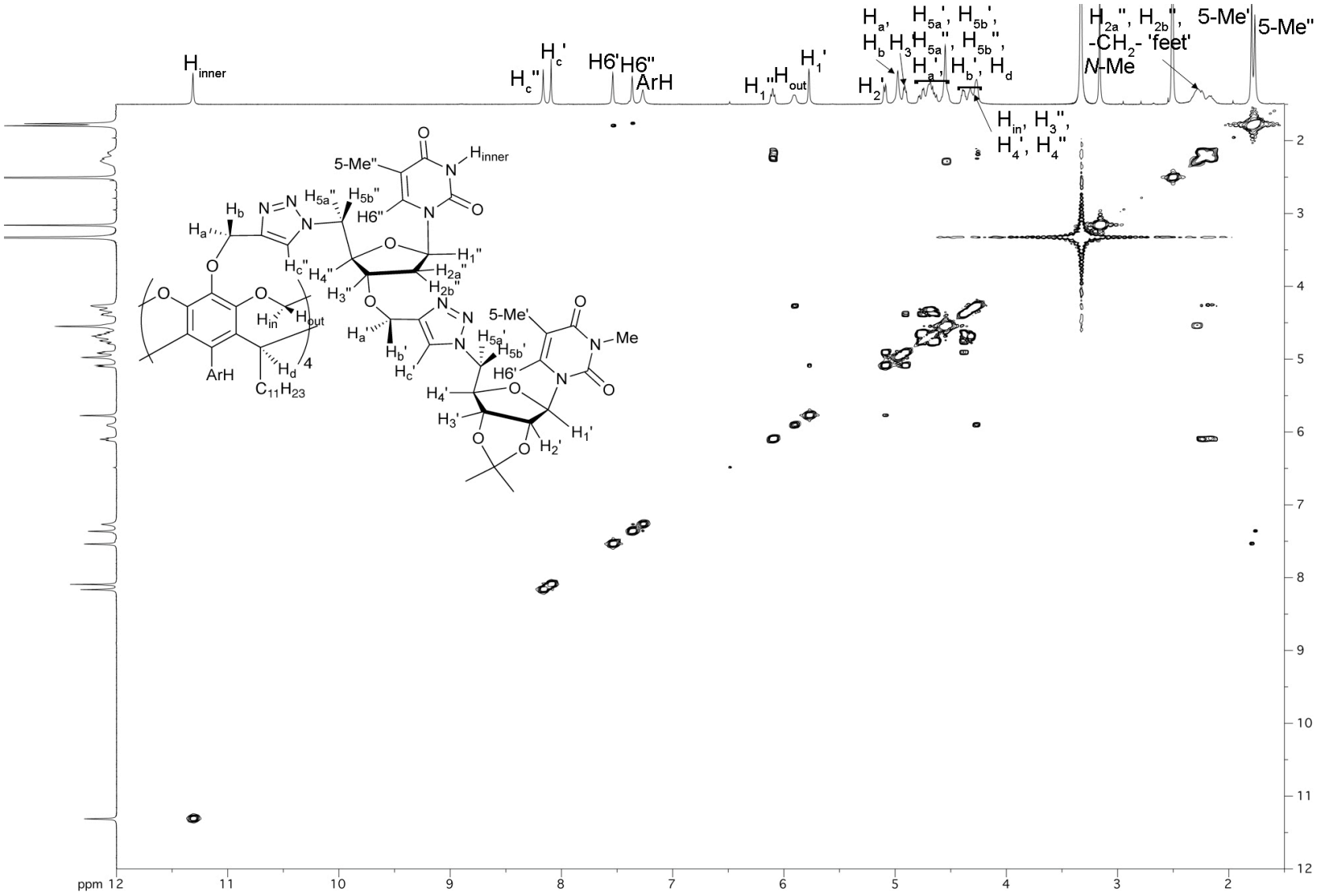


Figure 4.24 ^{1}H-^{1}H COSY spectrum of 47 in DMSO-d_6 at 25 ºC.

Chapter 5: Conclusion

5.1 Thesis summary

Nucleic acid topology is a derivative of the complex interweaving of structural motifs adopted by the constituent nucleobases, and is essential for normal cellular function. Directing this topology are intermolecular forces such as electrostatic interactions, hydrogen bonding and π stacking. G-quartets and quadruplexes are prime examples of tertiary nucleic acid structure, and their ubiquity in the telomeric end regions has implicated their involvement in such fundamental biological processes as replication, senescence and apoptosis. Pharmaceutical targeting of these structures has therefore been one of the foci of anticancer therapy. Construction of these motifs for *in vitro* model studies has been facilitated by organic template molecules that lower the entropic penalty of organizing multiple guanine subunits, thereby increasing thermodynamic stability. This approach has been used extensively by our group with the use of rigid, resorcinol-based cavitands to stabilize discrete G-quartets in both lipophilic and hydrophilic media. These template-assembled synthetic G-quartets (TASQs) have been shown to be viable substrates for drug binding assays.

Extension of these systems from bicyclic guanine-based to monocyclic pyrimidine-based assemblies was the driving force behind this thesis. Uracil and thymine-based quartets are rare due to their inherent instability, and when encountered, are always confined within G-quadruplex scaffolds and have been shown to augment the thermal stability of these structures. Corresponding U- and T-quadruplexes have not been reported until now. Further study of these systems *in vitro* might pave the way towards a greater understanding on their effects on overall nucleic acid stability and topology. As a major goal of this thesis, it was envisioned that discrete

U- and T-quartets and quadruplexes could be stabilized by cavitand templates in solution. Following this, these motifs were prepared by the conjugation of uridine and thymidine units to cavitands via the copper (I)-catalyzed azide-alkyne cycloaddition, commonly referred to as the "click" reaction.

Chapter 2 describes the synthesis and characterization of conjugate **1** that involved the installation of four uridine nucleosides on a lipophilic cavitand bearing pendant long-chain aliphatic 'feet'. Preliminary ^{1}H NMR spectroscopic analysis in $CDCl_3$ indicated involvement of the bound uracil bases in hydrogen bonding in contrast with the free nucleoside. NOESY experiments ascertained the assembly of a cation-free, all-*syn* U-quartet at room temperature, which becomes more tightly associated at –20 ºC. This was confirmed by variable temperature spectra observing a greater degree of hydrogen bonding with decreasing temperature. A planar U-quartet is proposed. DOSY NMR was instrumental in establishing the unimolecular state of the conjugate. Evidence for stacking of the uracil and underlying triazole rings via π-π interactions was also presented and believed to further stabilize the quartet. CD spectroscopic analysis indicated that the U-quartet is destabilized in a polar protic solvent such as methanol, a hydrogen bond disruptor.

Cationic extraction studies were also performed to investigate the affinity of the U-quartet towards metal cations such as K^+, Na^+ and Sr^{2+}, which are known to bind and stabilize guanine-based architectures. These cations were introduced in the form of their metal picrate salts. K^+ and Na^+ were found not to be taken up, while Sr^{2+} is sequestered by the conjugate to possibly form a symmetric homodimer in chloroform. ^{1}H and DOSY experiments provided indication of the bimolecular nature of the complex, while CD spectral traces observed its denaturation in methanol.

Chapter 3 details efforts towards the construction of an unprecedented U-quadruplex in the solution state as an extension of the work in the previous Chapter. The design, synthesis and characterization of conjugate **18**, composed of four uridine dinucleoside residues coupled to a cavitand template was completed. Preliminary ^{1}H NMR in $CDCl_3$ revealed a severely broadened spectrum indicative of higher order aggregation. NOESY elucidated the assembly of a cation-free, two-tiered U-quadruplex in DMSO-d_6 at room temperature with all uridine residues configured *syn*. Evidence that the quadruplex is closed to the extraneous solvent environment was also presented. Added stability is potentially derived from stacking of the uracil and triazole rings via π-π interactions. Further NOE evidence suggests a non-planar outer U-quartet. A planar inner quartet and 'dipped' non-planar outer quartet is proposed. Variable temperature studies demonstrated that the quadruplex is thermally stable up to about 80 ºC. DOSY measurements substantiated the unimolecular state of **18** in DMSO-d_6. Corresponding measurements in $CDCl_3$ indicated the tetrameric aggregation of the conjugate in this solvent, confirming the preliminary hypothesis.

Chapter 4 reports the assembly of a T-quartet and quadruplex to further reflect the templating ability of cavitands, especially when extended to more unstable systems. The synthesis and characterization of conjugates **33** and **34**, cavitands bound to four thymidine-based nucleosides and dinucleosides respectively, was completed. Signal broadening was observed for both compounds in $CDCl_3$ at room temperature. Variable temperature ^{1}H NMR of **33** observed the presence of two differentially hydrogen-bonded species at low temperature (–20 and –40 ºC), indicating conformational exchange to be in operation and the cause of the signal broadening at room temperature. NOESY at –40 ºC in $CDCl_3$ uncovered the presence of a stacked, all-*syn* T-quartet as the major species in addition to a non-quartet minor species. CD

spectroscopy further demonstrated the absence of assembly at room temperature. Treatment of **33** with $Sr(pic)_2$ observed no induced CD signal to suggest uptake of the Sr^{2+} cation and consequent structure assembly. **34** was found to assemble a cation-free, two-tiered T-quadruplex in DMSO-d_6 at room temperature. The inner quartet is proposed to adopt a propeller-twisted arrangement, while the outer quartet remains planar. DOSY indicated the unimolecular state of this quadruplex, while dimerization of the conjugate was found in $CDCl_3$. Evidence for a fully ring-stacked scaffold as the result of intercalation of all triazole and thymine rings was also found. The significant stabilizing role played by π-π interactions was qualified by NOESY studies on a hemimethylated conjugate.

The work presented in this thesis highlights the first reported examples of discrete, intramolecular uracil and thymine-based quartets and quadruplexes in the solution state, successfully addressing the main aim of the thesis. Characterization of these assemblies in organic media has been facilitated by the exceptional lipophilicity conferred by the long chain aliphatic C_{11} 'feet' of the cavitand templates. Most notably, the indispensable role of the templates in directing self-assembly and stabilizing these nucleobase scaffolds was demonstrated throughout this work.

5.2 Future work

5.2.1 Synthesis of cavitand 'foot'-based nucleobase conjugates

The basket-like geometry of cavitands gives rise to a upper 'rim' tapering into a narrower bottom 'foot' (Figure 5.1). Self-assembly at the 'foot' position has not been investigated in our group, with the exception of a T-quartet reported in polar protic solvent.[130] This narrower position should, in principle, be able to better stabilize tetranucleoside-based assemblies by positioning the self-assembling components in closer proximity.

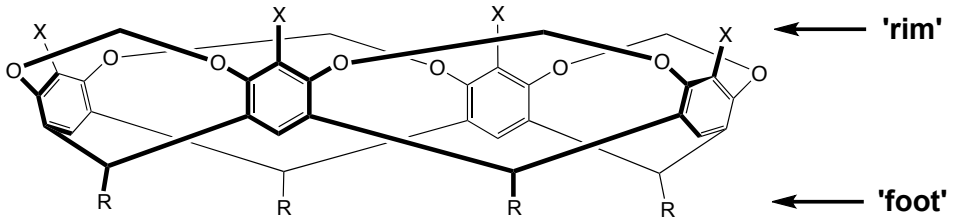
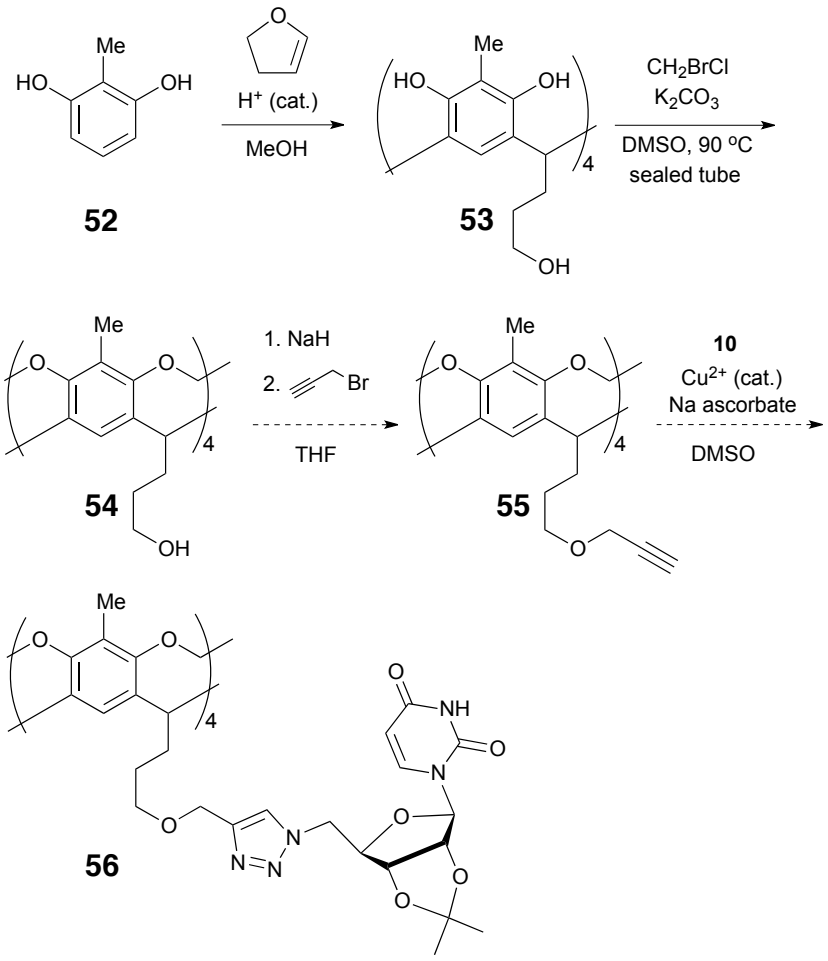


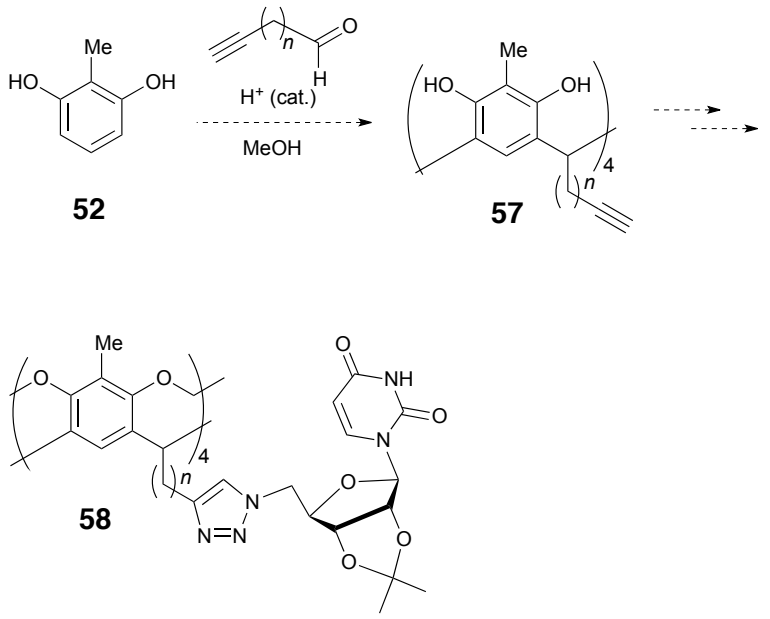
Figure 5.1 Side-on representation of a cavitand with rim and foot positions labeled.

Linker lengths could be easily modified in the first step of cavitand synthesis i.e. condensation of resorcinol and various aldehyde and masked aldehyde subunits to give the corresponding octols. For example, 2-methylresorcinol **52** can be condensed with 2,3-dihydrofuran to form the macrocyclic octol **53**, which on bridging provides cavitand **54**. This synthesis has been reported.[131] Subsequent deprotonation of the pendant alcoholic 'feet' with a strong base like sodium hydride and treatment with propargyl bromide could be effected to yield **55**. Coupling with the appropriate nucleosides or dinucleosides e.g. **10** can then proceed by the Cu-catalyzed "click" reaction to access the corresponding conjugates e.g. **56** (Scheme 5.1).



Scheme 5.1 Proposed synthesis of 'foot'-based tetrauridine-cavitand conjugate 56.

Alternatively, the condensation of terminal alkynyl aldehydes with **52** could provide cavitands of varying 'feet' lengths, allowing investigation of their effect on the stability of quartets and quadruplexes (Scheme 5.2).



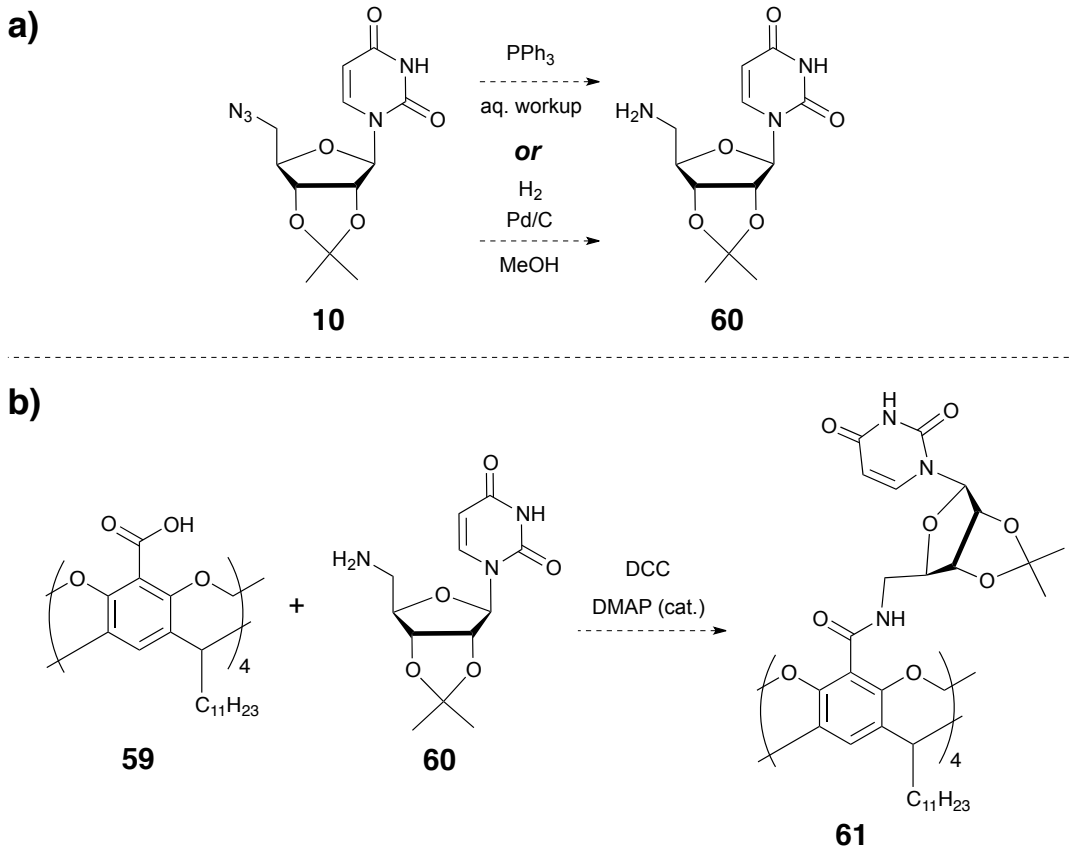
Scheme 5.2 Proposed synthesis of conjugates of varying 'feet' lengths.

5.2.2 Synthesis of amide-linked conjugates

The conjugation of nucleosides to cavitand templates with amide linkages is also envisioned. This concept is derived from peptide nucleic acids (PNA), synthetic analogs of DNA and RNA, whose backbones are comprised entirely of amide bonds.[132-135] PNA duplexes have been shown to have greater thermal stability than their DNA counterparts due to the lack of negatively charged phosphate groups and the resultant absence of electrostatic repulsion.[136] For this reason, PNA oligomers have been used as antisense markers for therapeutic applications.[137-139] They are also resistant to enzymatic degradation by both nucleases and proteases,[140] which are unable to recognize them as substrates.

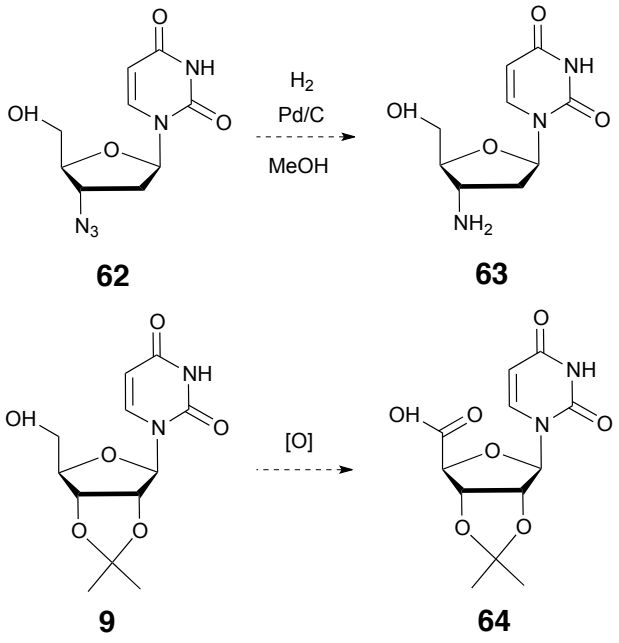
Use of the amide linker in nucleoside-cavitand conjugates may be advantageous for its smaller size compared to the triazole ring, and for having greater rotational freedom. The conjugation may be effected by a Staudinger ligation[141, 142] between an amine and carboxylic

acid in the presence of DCC coupling reagent and DMAP as catalyst. Cavitands such as **59** with carboxyl-functionalized rims have been reported. 5'-amino nucleosides e.g. **60** can be readily prepared by the reduction of the 5'-azido derivatives with triphenylphosphine in the Staudinger reduction, or with hydrogen gas over a palladium-carbon catalyst. Successful condensation of these two reactants affords the desired conjugate **61**. Scheme 5.3 shows the proposed synthesis of **60** and **61**.



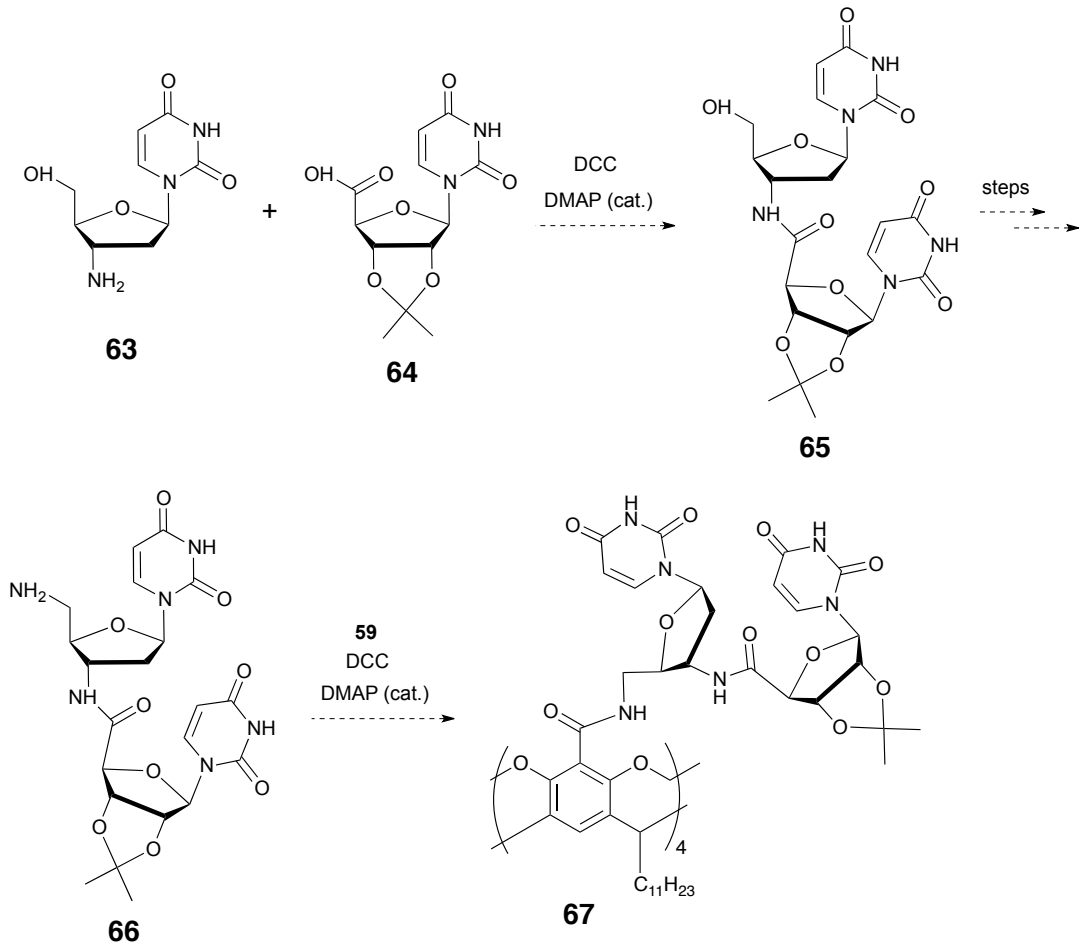
Scheme 5.3 a) Synthesis of 5'-amino nucleoside 60. b) DCC-mediated Staudinger ligation between cavitand 59 and nucleoside 60.

This methodology could also be translated to an amide-linked dinucleoside for the construction of a representative quadruplex system. The synthesis of the dinucleoside could begin with the coupling of derivatized nucleosides **63** and **64**. Preparation of **63** can be achieved with the reduction of the uracil analog of anti-retroviral drug 3'-azidothymidine (AZT) **62**, which can be prepared following a literature procedure,[143] while **64** can be accessed with the oxidation of **9** with standard chromium (VI) chemistry (Scheme 5.4) or several other methods.



Scheme 5.4 Proposed syntheses of nucleosides 63 and 64.

DCC coupling of **63** and **64** provides 5'-hydroxyl dinucleoside **65**. Functional group transformation to the 5'-amino group of **66** could be achieved over several steps. A final DCC coupling of **66** to cavitand **59** is then effected to afford conjugate **67** (Scheme 5.5).



Scheme 5.5 Proposed synthesis of conjugate 67.

Construction of longer length oligomers could then be pursued to function as potential model systems for various bioassays. Such systems may also find use as deep channel ionophores, or as DNA or RNA targeting vectors for antisense studies.

References

(1) Lehn, J.-M. *Science* **1993**, *260*, 1762 – 1763.

(2) Lehn, J.-M.; In *Supramolecular Chemistry: Concepts and Perspectives*; Wiley-VCH, 1995.

(3) Fyfe, M. C.; Stoddart, J. F. *Acc. Chem. Res.* **1997**, *30*, 393-401

(4) Seidel, S. R.; Stang, P. J. *Acc. Chem. Res.* **2002**, *35*, 972-983

(5) Correa-Duarte, M. A.; Perez-Juste, J.; Sanchez-Iglesias, Ana.; Glersig, M.; Liz-Marzan, L. M. *Angew. Chem. Int. Ed.* **2005**, *44*, 4375-4378

(6) Georgakilas, V.; Gournis, D.; Tzitzios, V.; Pasquato, L.; Guldi, D. M.; Prato, M. *J. Mater. Chem.* **2007**, *17*, 2679-2694

(7) Lee, C. H.; Tian, L.; Abbas, A.; Kattumenu, R.; Singamaneni, S. *Nanotechnology* **2007**, *22*, 275311

(8) Dickerson, M. B.; Sandhage, K. H.; Naik, R. R. *Chem. Rev.* **2008**, *108*, 4935-4978

(9) Singh, Y.; Dolphin, G. T.; Razkin, J.; Dumy, P. *ChemBioChem* **2006**, *7*, 1298-1314

(10) Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F. *Science* **2008**, *321*, 1795-1799

(11) Cram, D. J. *Science* **1983**, *219*, 1177-1183

(12) Gibb, B. C.; Mezo, A. R.; Causton, A. S.; Fraser, J. R.; Tsai, F. C. S.; Sherman, J. C. *Tetrahedron* **1995**, *51*, 8719-8732

(13) Mezo, A. R.; Sherman, J. C. *J. Am. Chem. Soc.* **1999**, *121*, 8983-8994

(14) Yang, H.; Sherman, J. C. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1752-1753

(15) Levene, P. *J Biol Chem* **1919**, *40*, 415-424

(16) Avery, O.; MacLeod, C.; McCarty M. *J Exp. Med.* **1944**, *79*, 137-158

(17) Hershey, A.; Chase, M. *J. Gen. Physiol.* **1952**, *36*, 39-56

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

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

(20) Chargaff, E.; Lipshitz, R.; Green, C. *J. Biol. Chem.* **1952**, *195*, 155-160

(21) Tsuchiya, Y.; Tamura, T.; Fujii, M. *J. Phys. Chem.* **1988**, *92*, 1760-1765

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

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

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

(25) 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

(26) 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

(27) Bang, I. *Biochem. Z.* **1910**, *26*, 293-311

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

(29) Chantot, J. F.; Guschlba, W. *FEBS Lett.* **1969**, *4*, 173-176

(30) Wlodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. *J. Phys. Chem. B* **2005**, *109*, 3594-3605

(31) Hud, N. V.; Smith F. W.; Anet, F. A. L.; Feigon, J. *Biochemistry* **1996**, *35*, 15383-15390

(32) Gu, J. ; Leszczynski, J. *J. Phys. Chem A*, **2002**, *106*, 529-532

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

(34) Zimmerman, S. B. *J. Mol. Bio.* **1976**, *106*, 663-672

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

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

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

(38) Sen, D.; Gilbert, W. *Nature* **1990**, *344*, 410-414

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

(40) Olovniko, A. *J. Theor. Biol.* **1973**, *41*, 181-190

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

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

(43) Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. *Genes Dev*. **1997**, *11*, 2801-2809

(44) Cuesta, J.; Read, M. A.; Neidle, S. *Mini Rev. Med. Chem*. **2003**, *3*, 11-21

(45) Neidle, S.; Read, M. A. *Biopolymers* **2000/2001**, *56*, 195-208

(46) Kerwin, S. M. *Curr. Pharm. Des.* **2000**, *6*, 441-478

(47) Cairns, D.; Anderson, R. J.; Perry, P. J.; Jenkins, T. C. *Curr. Pharm. Des.* **2002**, *8*, 2491-2504

(48) Li, J.-L.; Harrison, R. J.; Reszka, A. P.; Brosh Jr., R. M.; Bohr, V. A.; Neidle, S.; Hickson, I. D. *Biochemistry* **2001**, *40*, 15194-15202

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

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

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

(52) 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

(53) Meyne, J.; Ratliff, R. L.; Moyzis, R. K. *Proc. Natl. Acad. Sci. U. S. A*. **1989**, *86*, 7049-7053

(54) Feng, J. L.; Funk, W. D.; Wang, S. S.; Weinrich, S. L.; Avilion, A. A.; Chiu, C. P.; Adams, R. R.; Chang, E.; Allsopp, R. C.; Yu, J.; et al. *Science* **1995**, *269*, 1236-1241

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

(56) Smithfeigon, F. W. J. *Nature* **1992**, *356*, 164-168

(57) Smith, F. W.; Schultze, P.; Feigon, J. *Structure* **1995**, *3*, 997-1008

(58) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Bio.* **2002**, *320*, 189-200

(59) Laughlan, G.; Murchie, A. I.; Norman, D. G.; Moore, M. H.; Moody, P. C.; Lilley, D. M.; Luisi, B. *Science* **1994**, *265*, 520-524

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

(61) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876-880

(62) Langridge, R.; Rich, A. *Nature* **1963**, *198*, 725-728

(63) Marsh, R. E.; Bierstedt, R.; Eichhorn, E. L. *Acta. Crystallogr.* **1962**, *15*, 310-316

(64) Akrnimiski, E. O.; Sander, C.; Ts'o, P. O. P. *Biochemistry* **1963**, *2*, 340-344

(65) Hartman Jr., J. A.; Rich, A. *J. Am. Chem. Soc.* **1965**, *87*, 2033-2039

(66) Inman, R. B. *J. Mol. Bio.* **1964**, *9*, 624-637

(67) Gehring, K.; Leroy, J.-L.; Gueron, M. *Nature* **1993**, *363*, 561-565

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

(69) Green, D. W.; Mathews, F. S.; Rich, A. *J. Biol. Chem*. **1962**, *237*, 3573-3575

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

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

(72) Phan, A. T.; Gueron, M.; Leroy, J.-L. *J. Mol. Biol.* **2000**, *299*, 123-144

(73) Phan, A. T.; Leroy, J.-L. *J. Biomol. Struct. Dyn.* **2000**, *17*, 245-251

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

(75) Davis, J. T.; Spada, G. P. *Chem. Soc. Rev*. **2007**, *36*, 296-313

(76) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A. *Angew. Chem. Int. Ed.* **2000**, *112*, 1356-1359

(77) 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

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

(79) Nikan, M.; Sherman, J. C. *J. Org. Chem*. **2009**, *74*, 5211-6218

(80) Stefan, L.; Guedin, A.; Amrane, S.; Smith, N.; Denat, F.; Mergny, J.-L.; Monchaud, D. *Chem. Commun.* **2011**, *47*, 4992-4994

(81) Xu, H. J.; Stefan, L.; Haudecoeur, R.; Vuong, S.; Richard, P.; Denat, F.; Barbe, J. M.; Gros, C. P.; Monchaud, D. *Org. Biomol. Chem*. **2012**, *10*, 5212-5218

(82) Marlow, A. L.; Mezzina, E.; Spada, G. P.; Masiero, S.; Davis, J. T.; Gottarelli, G. *J. Org. Chem*. **1999**, *64*, 5116-5123

(83) Forman, S. L.; Fettinger, J. C.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067

(84) Shi, X. D.; Fettinger, J. C.; Davis, J. T. *J. Am. Chem. Soc.* **2001**, *123*, 6738-6739

(85) Shi, X. D.; Fettinger, J. C.; Davis, J. T. *Angew. Chem. Int. Ed.* **2001**, *40*, 2827-2831

(86) Nikan, M.; Patrick, B. O.; Sherman, J. C. *ChemBioChem* **2012**, *13*, 1413-1415

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

(88) Zhao, X.; Meng, F. *J. Mol. Struct.: THEOCHEM* **2006**, *770*, 157-162

(89) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2001**, *105*, 10366-10371

(90) Meyer, M.; Steinke, T.; Brandl, M.; Suhnel, J. *J. Comput. Chem.* **2001**, *22*, 109-124

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

(92) Barcelo-Oliver, M.; Estarellas, C.; Terron, A.; Garcia-Raso, A.; Frontera, A. *Chem. Commun.* **2011**, *47*, 4646-4648

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

(94) Pan, B.; Xiong, Y.; Shi, K.; Sundaralingam, M. *Structure* **2003**, *11*, 825-831

(95) Yan, X.; Ishizuka, T.; Kimura, T.; Komiyama, M. *J. Am. Chem. Soc.* **2010**, *132*, 7231-7233

(96) Patel, P. K.; Hosur, R. V. *Nucleic Acids Res.* **1999**, *27*, 2457-2464

(97) Cacares, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097-1102

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

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

(100) Roman, E.; Peinador, C.; Mendoza, S.; Kaifer, A. E. *J. Org. Chem.* **1999**, *64*, 2577-2578

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

(102) Rostovtsev, V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599

(103) Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V.; Noodleman, L.; Sharpless, K. B. *J. Am. Chem. Soc.* **2005**, *127*, 210-216

(104) Huisgen, R. *Angew. Chem. Int. Ed.* **1963**, *2*, 633-645

(105) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11-15

(106) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870-2871

(107) Waldeck, A. R.; Kuchel, P. W.; Lennon, A. J.; Chapman, B. E. *Prog. Nucl. Magn. Reson.* **1997**, *30*, 39-68

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

(109) Price, W. S. *Concepts Magn. Reson.* **1997**, *9*, 299-336

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

(111) Evan-Salem, T.; Baruch, I.; Avram, L.; Cohen, Y.; Palmer, L. C.; Rebek, J. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12296-12300

(112) Avram, L.; Cohen, Y. *J. Am. Chem. Soc.* **2002**, *124*, 15148-15149

(113) Timmerman, P.; Weidmann, J. L.; Jolliffe, K. A.; Prins, L. J.; Reinhoudt, D. N.; Shinkai, S.; Frish, L.; Cohen, Y. *J. Chem. Soc., Perkin Trans. 2* **2000**, 2077-2089

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

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

(116) Clark, L. B. *J. Am. Chem. Soc.* **1994**, *116*, 5265-5270

(117) Fulsher, M. P.; Serrano-Andres, L.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 6168-6176

(118) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061-4067

(119) Gray, D. M.; Bollum, F. J. *Biopolymers* **1974**, *13*, 2087-2102

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

(121) Miles, D. W.; Robins, M. J.; Robins, R. K.; Winkley, M. W.; Eyring, H. *J. Am. Chem. Soc.* **1969**, *91*, 824-831

(122) Bureekaew, S.; Hasegawa, J.; Nakatsuji, H. *Chem. Phy. Lett.* **2006**, *425*, 367-371

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

(124) Lucas, R.; Neto, V.; Bouazza, A. H.; Zerrouki, R.; Granet, R.; Krausz, P.; Champavier, Y. *Tetrahedron Lett.* **2008**, *49*, 1004-1007

(125) McGee, D. P. C.; Martin, J. C. *Can. J. Chem*. **1986**, *64*, 1885-1889

(126) Wang, S. M.; Li, N. C. *J. Am. Chem. Soc.* **1968**, *90*, 5069-5074

(127) Schweighart, F. K.; Moll, C.; Li, N. C. *J. Mag. Res.* **1970**, *2*, 35-41

(128) Dyllick-Brenzinger, C.; Sullivan, G. R.; Pang, P. P.; Roberts, J. D. *Proc. Natl. Acad. Sci. U. S. A*. **1980**, *77*, 5580-5582

(129) Petersen, S. B.; Led, J. J. *J. Am. Chem. Soc.* **1981**, *103*, 5308-5313

(130) Bare, G. A. L.; Sherman, J. C. *J. Org. Chem*. **2013**, *78*, 8198-8202

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

(132) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500

(133) Corey, D. R. *Trends Biotechnol.* **1997**, *15*, 224-229

(134) Nielsen, P. E.; Egholm, M. *Curr. Issues Mol. Biol.* **1999**, *1*, 89-104

(135) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897

(136) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 2796-2823

(137) Demidov, V. V. *Drug Discov. Today* **2002**, *7*, 153-155

(138) Nielsen, P. E. *Curr. Opin. Struct. Biol.* **1999**, *9*, 353-357

(139) Henvey, J. C.; Peffer, N. J.; Bisi, J. E. *Science* **1992**, *258*, 1481-1485

(140) Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sonnichsen, S. H.; et al. *Biochem. Pharmacol.* **1994**, *48*, 1310-1313

(141) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010

(142) Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 2686-2695

(143) Czernecki, S.; Valery, J.-M. *Synthesis* **1991**, 239-240