Synthesis of novel molecules for specific recognition of DNA G-quadruplex structures and rapid analysis of their binding to telomeric DNA on a gold surface.

Killian Walsh, B.Sc. (Hons.)

Under the supervision of:
Prof Richard O`Kennedy, School of Biotechnology, DCU,
Dr Vladimir Gubala, Medway School of Pharmacy, University of Kent,
and Dr Kieran Nolan, School of Chemical Sciences, DCU.

A thesis presented to Dublin City University
for the degree of Master of Science

School of Biotechnology,
Dublin City University,
September 2014.

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Master of Science is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Killian Walsh ID No.: 56072941 Date:
Signed: _____

Contents

Abbreviations

(NMMIX)	N-methyl mesoporphyrin IX
8ArGdG	8-Aryl-2'-deoxyguanosine analogues
8PhG	8-Phenyl-2'-deoxyguanosine analogues
BAIB	Bisacetoxyiodobenzene
BDI	bio diagnostics institute
BRACO-19	9-[4-(N,N-dimethylamino)phenylamino]-3,6-bis(3-pyrrolodino-propionamido)-acridine
CBs	Cajal bodies
CGLT	Covalently linked G-tetrad
COSY	Correlation spectroscopy
COSY	Correlation Spectroscopy
DCM	dichloromethane
dG	2'-deoxyguanosine
DMAP	4-Dimethylaminopyridine
DMF	dimethylformamide
GN	Guanine nucleotides
GQ	Guanine -Quadruplex
GRS UTRdb	Guanine -rich sequence untranslated region database
GRSDB	The G-rich Sequences Database
GRSDB2	The G-rich Sequences Database 2
HMBC	Heteronuclear Multiple Bond Correlation
HMQC	Heteronuclear Multiple Quantum Correlation
HPLC	High performance liquid chromatography
MeOH	methanol
NOESY	Nuclear Overhauser Enhancement Spectroscopy
PBS	phosphate buffered saline solution
PDC	micron Piezo Dispense Capillary
POT1	Protection of telomeres protein 1
RAP1	Telomeric repeat-binding factor 2-interacting protein 1
RDP	Ribosomal Database Project
RHPS4	3,11-Difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl]acridinium methosulfate
RR82	Pyridostatin
Sulfo-SMCC	Sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate

TCAB1	Telomerase Cajal body protein 1
TCEP	(tris(2-carboxyethyl)phosphine)
T-dGTP	6-thio-7-deaza-2'-deoxyguanosine 5'-triphosphate
TEA	Triethylamine
TEMPO	2,2,6,6-Tetramethylpiperidinyloxy
TERT	Telomerase Reverse Transcriptase
TERT	telomerase reverse transcriptase
THF	tetrahydrofuran
TIN2	Telomeric repeat binding- factor 1-interacting nuclear factor 2
TLC	Thin Layer Chromatography
T-Loop	Telomere-Loop
TPP	Triphenylphosphine
TPP1	Adrenocortical dysplasia protein homolog
TPPTS	3,3′,3″-Phosphanetriyltris(benzenesulfonic acid) trisodium salt
TR	Telomerase template RNA
TRF1	Telomeric repeat binding- factor 1
TRF2	Telomeric repeat binding- factor 2

Abstract

When confronted with nucleobase interactions, most people will think of the Watson-Crick base pairs. However, research from a variety of different disciplines has shown that the use of nucleobases as a scaffold to build larger functional assemblies through a variety of noncovalent interactions is much more versatile. Although all nucleobases have the ability to homodimerize, homo-oligomers of guanine derivatives are perhaps the most interesting. One of the most distinctive examples is the G-tetrad, a hydrogen-bonded macrocycle formed by cation-templated assembly of guanosine. In the late 1980s it was suggested that this motif, when formed in DNA, might be biologically relevant, which sparked enormous research interest in these naturally occuring architectures.

The scope of the present work also relates to the G-quartet motif, and it fits particularly well with contemporary studies in molecular self-assembly and noncovalent synthesis. A series of 8-aryl-2'-deoxyguanosine (8ArdG) analogues was synthesised and their ability to induce G-quadruplex formation was investigated. Four lead compounds were selected to analyse the variation in G-quadruplex stability based on the position of functional groups. The detection method used was of novel design, based on quenching of immobilised fluorescently labelled DNA on a gold surface. The DNA undergoes structural changes upon binding to the 8ArdG, indicating the formation of G-quadruplex, which can be detected as a decrease in fluorescence signal. The formation of G-quadruplex was further validated by the use of Dual Polarisation Interferometry (DPI). A library of three novel compounds a one previously reported compound was successfully synthesised and fully characterised by NMR and IR spectroscopy. One of the novel compounds, 8-(4-pyrazole)-2`-deoxyguanosine, displayed novel characteristics as revealed by two dimensional NMR analysis. This novel compound displayed unusual chemical shifts attributed to a previously unreported molecular conformation. DPI analysis, using thiolated telomere DNA successfully bound to surface of amine coated chip, has revealed self-assembly in the presence of 8-(4-pyrazole)-2`-deoxyguanosine. As it is shown in this work, the noncovalent bond is a powerful synthetic tool in the preparation of complex molecular architectures. The G-quartet in particular serves as an excellent model to learn about the fundamentals and utility of noncovalent interactions in molecular self-assembly and drug discovery.

Chapter 1

Introduction

1.1 Cell Replication

The process of cell replication begins with duplication of DNA, which is an essential process for all eukaryotic life forms and allows for the proliferation of all cells. DNA replication produces a nearly perfect duplicate of the original DNA strand. However, as DNA replication continues throughout the life-cycle of a cell, errors and breaks occur in the genetic sequence and the replicates continually degrade from the original DNA genetic template. In order to prevent and ultimately minimise these errors in replication, eukaryotic cells utilise cellular proofreading and repair mechanisms that ensure near perfect fidelity for DNA replication. Although sophisticated, the cell's DNA repair mechanisms cannot distinguish between breaks in the DNA sequence and the natural ends of chromosomes. Interaction of repair mechanisms and the chromosome terminals would lead to adverse chromosome fusions and, therefore, have to be avoided[1]**.** To facilitate this goal the cell has evolved a protective "cap" on the strands of exposed chromosomes called a telomere.

1.2 Telomere

A telomere is a region of repetitive nucleotides found at the ends of chromosomes in the sequence TTAGGG.[2] In eukaryotic cells, DNA replication enzymes (DNA polymerase protein complex) cannot replicate the sequences present at the ends of the chromosomes. Hence, these sequences and the information they carry may get lost after each cell replication. The purpose of the telomere is to protect the chromosome, by serving as a disposable buffer which safeguards the chromosome from the loss of genetic information during cell replication. If the telomere was to be completely depleted, the DNA sequences in the cells chromosomes would be left unprotected and susceptible to the loss of vital genetic information. Any additional replication by such cells could lead to the formation of defective cells, the effects of which could manifest as genetic diseases or death for the host body. A telomere is comprises double-stranded and single-stranded DNA. Where the telomere diverges from double to single strand, there exist two separate terminal ends, the 5`-terminus and the 3`-terminus[3], as illustrated in (Figure 1.1).

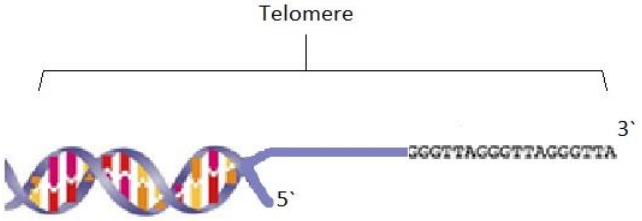


Figure 1.1: *A cartoon depicting the 3`-and 5`-termini of a telomere strand comprised of a TTAGGG sequence of base pairs*[4]*.*

The process by which the telomere 3` and 5` distinguish themselves from DNA breaks to the DNA repair machinery is currently not fully understood; however, the shape and orientation of the telomere strand has been proposed to play a pivotal role in this regard. The actual end of the telomere is a protrusion of single stranded repeats from the 3' end, termed the G-tail or G-overhang. It has been proposed that the 3' G-overhang is a 'hairpin-like structure known as the Telomere-Loop (T-Loop). This is shown in Figure 1.2. The T-Loop orients itself into a folded shape that shields the 3`OH group from exposure to DNA repair mechanisms.

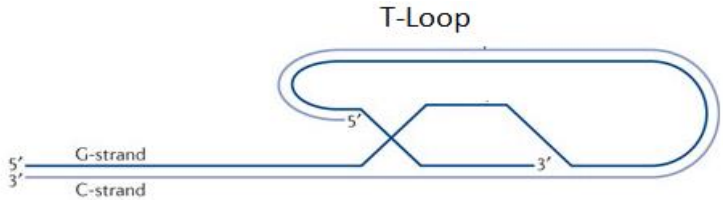


Figure 1.2: *Telomere Loop (T-loop) commonly referred to as a G-tail or G-overhang, is a 'hairpin-like' structure adopted by the terminal ends of telomeres*[5].

1.3 Telomerase Reverse Transcriptase

When a cell replicates, the cell`s telomere strand is truncated. This process limits the number of times a cell can replicate, as continued replication would cause telomere strands to become critically short, resulting in cell senescence or apoptosis. Human cell populations can divide 40 to 60 times, then the cell reaches the Hayflick limit[6] and cell division stops. In order to survive, a eukaryotic organism`s cells require the ability to replicate beyond its Hayflick limit. To facilitate this goal, an enzyme called Telomerase Reverse Transcriptase (TERT) periodically replenishes the telomere strand Figure 1.3. The TERT enzyme is part of a larger telomerase complex, consisting of TERT, the RNA subunit TR (Telomerase template RNA) that provides the template for repeated telomere synthesis and Dyskerin, a key auxillary protein[7].

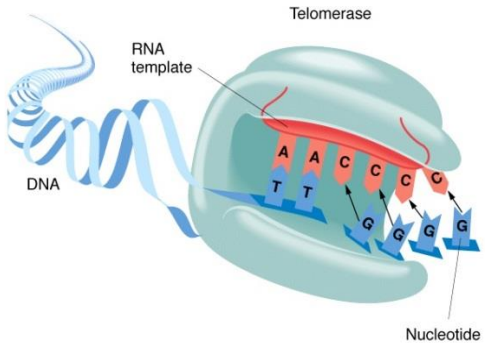


Figure 1.3: *The telomerase complex and the mechanism of action showing the RNA template (red) and the addition of complementary bases to elongate the telomere*[8].

The telomerase complex exists first as an immature complex, assembled in the nucleus of the cell, in sub-organelles called Cajal bodies (CBs). CBs are small subnuclear membraneless organelles present either free in the nucleoplasm and/or physically associated to specific regions

of chromatin. CBs contain newly assembled small nuclear ribonucleoproteins and help facilitate ribosomal RNA processing. The immature telomerase complex is shuttled to telomeres via TCAB1 (Telomerase Cajal body protein 1)[9]**.** The immature complex is sequestered into an active conformation by the ATPases, Pontin and Reptin, where it associates with the terminal exposed 3` hydroxyl group of the T-Loop[10]. Telomere length and integrity in humans is regulated by Shelterin, figure 1.4. Shelterin is a complex of six specific proteins, TRF1, TRF2, TIN2, Rap1, TPP1, and POT1 that distinguish and protect the telomere`s exposed 5` terminal from DNA repair machinery[11]. Shelterin is essential for the extension of telomeric DNA, as it aids the ability of telomerase to distinguish the 3` and 5` terminals. If telomerase were to extend the 5` terminal in place of the 3`, the chromosome ends come become fused, which would be detrimental for the cell.

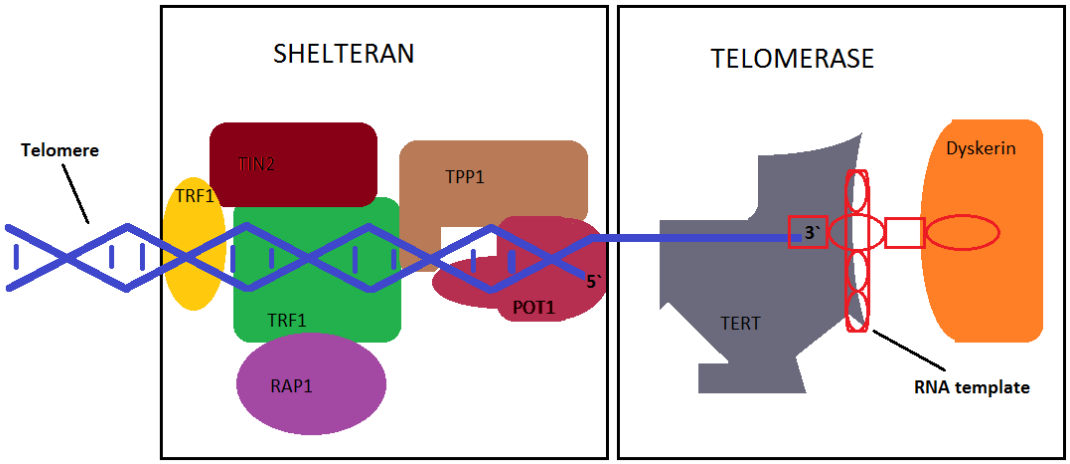


Figure 1.4: *The Telomerase-Shelterin complex consisting of Telomeric repeat binding-factor 1(TRF1), Telomeric repeat binding- factor 2 (TRF2), Protection of telomeres protein 1 (POT1,) Telomeric repeat-binding factor 2-interacting protein 1 (RAP1), Adrenocortical dysplasia protein homolog (TPP1), Telomeric repeat binding- factor 1-interacting nuclear factor 2(TIN2), protecting the 5` telomeric terminal overhang from interaction with telomerase reverse transcriptase (TERT), the Dyskerin protein localises to dense fibrillar components of nucleoli and to coiled Cajal bodies in the nucleus*[12].

1.4 Telomerase, Cancer and G-quadruplex

Although TERT is essential for the proliferation and survival of eukaryotic mammalian cells, overexpression/reactivation of TERT in cells can result in telomere strands being extended indefinitely. Cancerous cells undergo uncontrolled replication far beyond the Hayflick limit; thus, cancerous cells are often referred to as “immortal cells”. When cancerous cells continue to replicate unabated they lead to the formation of tumours inside the host body. Experiments by Kim *et al.* of cultured cells of eighteen different cancerous human tissues, resulted in 98 of 100 immortal cancerous cells and zero of 22 healthy non-cancerous cell populations testing positive

for telomerase[13]. On further testing human tumours tested positive for TERT in 90 of 101 biopsies, in contrast 50 samples of normal somatic tissues tested negative for TERT. The results also showed that benign tumours "such as fibroids" were negative. The conclusion drawn from these results was that TERT was repressed in healthy non-cancerous cells but reactivated in cancerous cells. This data was strong evidence of a correlation between the ability of cancerous immortal cells to continually replicate and the overexpression/reactivation of TERT.

1.5 Telomeric DNA and aging

The attrition of telomeres is thought to play an important role in cell deterioration with advancing age and subsequently the longevity of eukaryotic organisms. A series of studies on a variety of eukaryotes has demonstrated a trend that organisms with longer telomere strands have longer life-spans than organisms of the same species with shorter strands. Studies of this relationship have been hampered by the time-scale required to repeatedly measure the telomere lengths of test subjects, particularly in long-lived species, where life-span variation is greatest.
A study published in 2012 by Monaghan *et al.*[14] on zebra finches provided an interesting insight. The experiment was performed over a nine year period with a sample size of 99 test subjects. The telomere strands were examined repeatedly throughout the life-span of each test subject. The study showed that test subjects that displayed the longest telomeres had longer life-spans than those with shorter strands.

It has also been well established *in vitro* that the artificial extension of telomeres via the temporary reactivation of TERT has successfully allowed human cells to replicate beyond their Hayflick limit and become immortal cells[15]. Further *in vivo* experiments on mouse models[16] have demonstrated that reactivation of TERT and the subsequent telomere lengthening has successfully reversed the aging process of such mice. A study performed on an invertebrate worm species produced similar results[17]. To date no such experiments of artificial telomere extension have been performed *in vivo* on humans. In light of this data it has been postulated that the extension of telomere strands through temporary reactivation of TERT may be a potential means of decelerating the aging process.

1.6 G-quadruplexes

In 2009 Blackburn, Greider and Szostak were awarded the Nobel Prize in Medicine for the discovery of the mechanism of telomeric protection of chromosomes and the enzyme telomerase. In 1987 they discovered that guanine molecules could self-assemble into intramolecular structures called G-Quadruplexes (GQs) and proposed that these structures were

important for telomere function and might have been biologically relevant to mammalian cell replication[18].

Since the interaction between cancer and GQ formation in telomeres was published by Blackburn *et al.* a large body of research has been published in this field with a high emphasis on using small molecules as a targeted approach in the development of novel chemotherapeutics. In 1997 Hurley *et al.* demonstrated the inhibition of telomerase via G-quadruplex formation induced by a small molecule[19]. In the past decade there have been several reports of compounds that inhibit TERT via the formation of GQs in the presence of telomeric DNA figure 1.5.

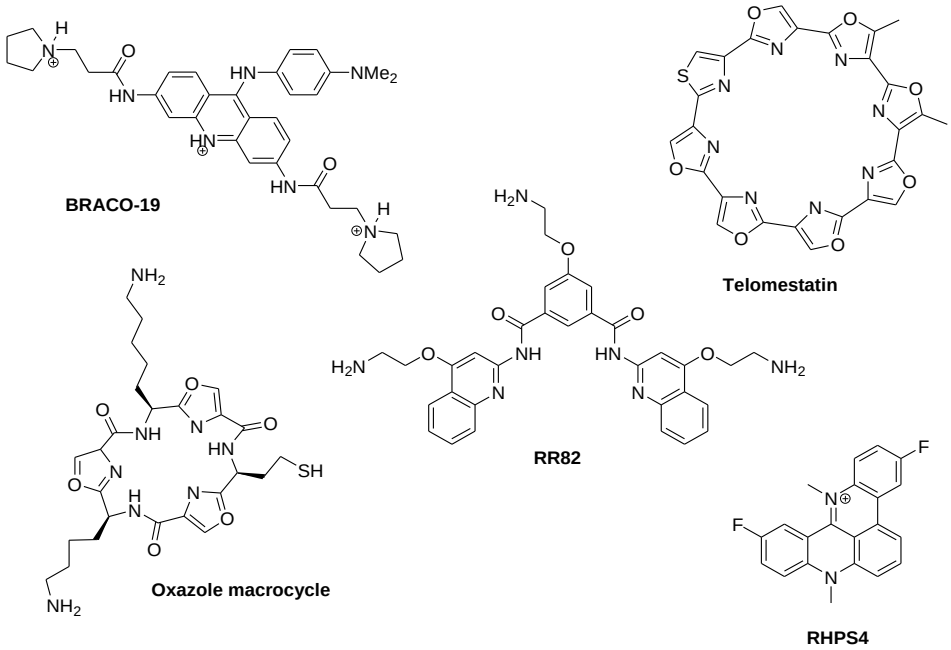


Figure 1.5: *Structures of quadruplex-forming ligands for which in vivo data have been reported for telomerase inhibition. BRACO-19, made from an acridine building block and modified in the 3, 6 and 9 positions. Telomestatin, is extracted from sea sponges. Pyridostatin (RR82) is synthesised from a pyridinedicarboxamide starting material and functionalized with 2 quinolinyl functional groups. RHPS4 is synthesised from cyclic polypeptide starting materials. The oxazole macrocycle was synthesised from peptide macrocycles and was modelled on the structure of telomestatin[20-23].*

G-quadruplex assemblies have been shown to form *in vitro* in a variety of organisms, e.g. in prokaryotic species such as bacteria[24] and in eukaryotic organisms such as plants[25] and in a variety of higher order vertebrates[26]. GQ assembly in eukaryotes *in vivo* has also been reported[27]. The human telomeric repeat is the same for all vertebrates. The GQs formed from this structure have been well studied by circular dichroism[28], ultraviolet absorbance[29], mass spectrometry[30], X-ray crystal structure determination[31] and Nuclear Magnetic Resonance spectroscopy (NMR)[32]. It has also been observed that GQs may also form in ‘guanine-rich’

strands of DNA other than the telomere single strand region located at chromosome terminal regions[33].

It is now well established that nucleic acids can undergo guanine-cytosine and adenine-thymine base pairing to form double helical structures such as DNA. However, guanine-based nucleotides such as deoxyguanosine can also self-assemble *via* Hoogsteen hydrogen bonding to adopt planar conformations of four G-subunits known as G-tetrads. Two or more G-tetrads can in turn assemble into higher order species *via* π-π stacking into a structure known collectively as a G-quadruplex. The quadruplex structure is further stabilised by the presence of a metal cation, such as potassium, which sits in a central channel between each pair of tetrads[34] figure 1.6.

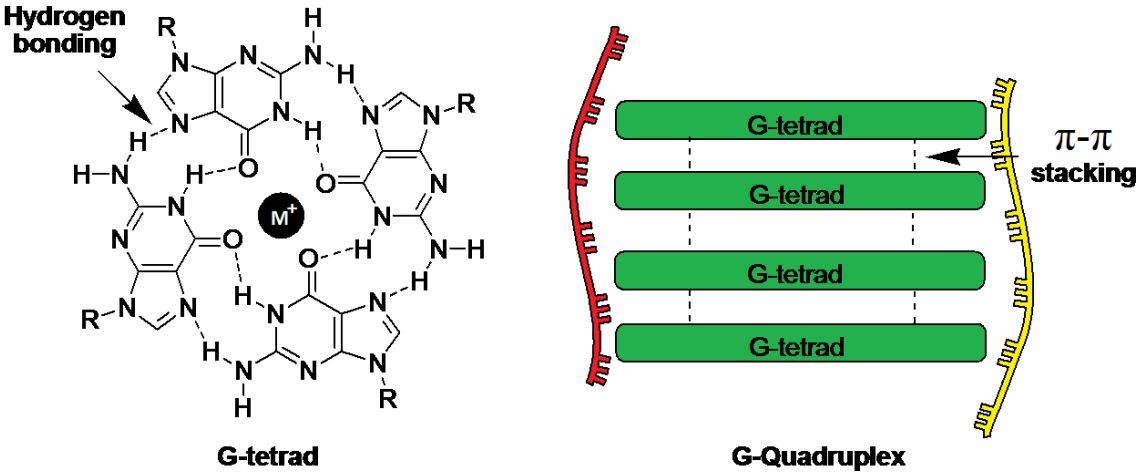


Figure 1.6: *G-tetrad and G-quadruplex arrangements. Four guanine subunits self-assemble via hydrogen bonding and then undergo π-π stacking to assemble into the G-quadruplex structure.*

G-quadruplexes are characterised by the vertical stacking of G-tetrads, through non-covalent interactions. The assembly begins with Hoogsteen hydrogen-bonding between each guanine unit into a cyclic G-tetrad arrangement. The G-tetrad is stabilised by a total of 8 hydrogen bonds between the (N1 to O6) and (N2 to N7)[35] as shown in Figure 1.6. DNA GQs are highly polymorphic and their structures are influenced by numerous factors such as temperature, pH, the DNA sequence and the ions present in solution. GQs are identified and distinguished according to their basic geometrical organisation.

GQs can assemble in an intra- or intermolecular fashion, and oligonucleotides containing one, two or four G-stretches can form tetrameric, dimeric or monomeric GQs, respectively. GQ formation in duplex DNA G-tracts can adopt parallel or anti-parallel orientations with respect to each other. Guanine derivatives can adopt syn or anti-glycosidic bond conformations. The loops connecting G-tetrads in GQs possess variable length and topologies, depending on the

directionality of the G-tracts that they connect. GQ topologies are defined by the combination of strand directionality, loop connectivity and groove dimensions with widths and depths determined by the overall geometrical factors described in figure 1.7 below[36].

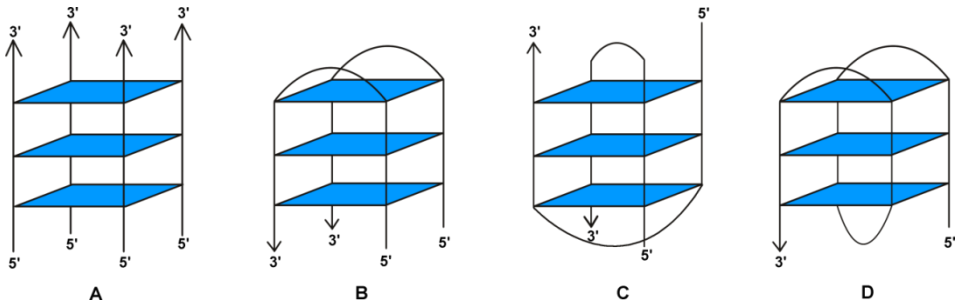


Figure 1.7: *The G-tail loop can fold in four different orientations to produce geometrically varying GQs. Parallel orientation of the loops will form a parallel stranded tetraplex. If the loops are folded into a hairpin dimerization conformation they can assemble in two ways. The two loops can assemble "edgewise" to form B, or if the loops assemble in a diagonally they will form C. Alternatively, a single loop can fold upon itself to form D (a unimolecular G-quadruplex).*

The topology of a GQ is also dictated by its method of assembly, which typically occurs in two different interactions as shown in figure 1.8. The first interaction is end-stacking, where the tetrads assemble one on top of another in an ordered sequence. The second interaction, intercalation, is when a guanine subunit inserts itself between G-tetrads. Data from reported quadruplex-ligand structures, obtained from NMR spectroscopy[37], fibre diffraction[38], computational chemistry and single-crystal X-ray crystallographic analysis[39,40] favour the end-stacking interaction.

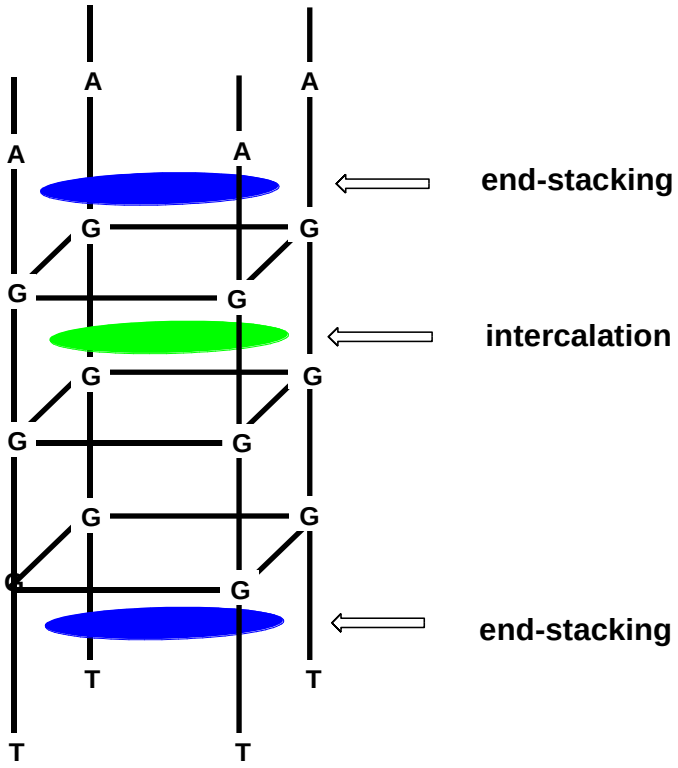


Figure 1.8: *Proposed interactions of planar aromatic drugs with GQs, depicting their end-stacking and intercalating positions.*

The specific ordered sequence of a GQ species has been demonstrated to have a high impact on the propensity of a molecule to bind selectively with G-tracts. This was demonstrated recently by Nicoludis *et al.*, where the molecule N-methyl mesoporphyrin IX (NMM) showed extremely high selectivity for endcap-binding to parallel GQs, whereas the same molecule demonstrated extremely poor selectivity for binding to non-parallel GQs[41].

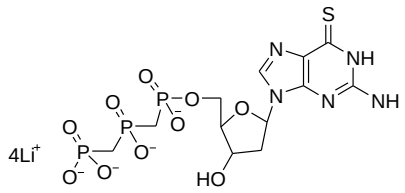
1.7 Deoxyguanosine: a model for potential chemotherapeutics

The telomeric DNA sequence is typically 'guanine-rich', comprised of a repeating motif of TTAGGG nucleic acids. Deoxyguanosine is the only nucleic acid that is known to self-assemble *in vitro* into GQ higher order structures. Since aromatic stacking is one of the main non-covalent interactions in the GQ DNA formation, an extended aromatic surface of any potential candidate molecule is required[42].

The overall size and three dimensional structure of deoxyguanosine make it a suitable model, on which to base a potential drug. It also possesses a relatively low molecular weight which would be beneficial if it can easily diffuse through biological membranes in order to interact at the

intended target site. Such criteria demonstrate that deoxyguanosine is an excellent scaffold on which to model potential novel derivatives to serve as chemotherapeutics *via* QG formation.

It has been demonstrated *in vitro* that derivatives of deoxyguanosine can inhibit the ability of TERT to extend telomeric DNA in human cells[43,44]. One such derivative, 6-thio-7-deaza-2'-deoxyguanosine 5'-triphosphate (T-dGTP), figure 1.9, is a potent inhibitor of human telomerase with an IC(50) of 60 nM. In light of the *in vitro* effects of various derivatives of deoxyguanosine to effectively inhibit TERT, it was postulated that cancerous cells could be selectively targeted by novel derivatives of deoxyguanosine as chemotherapeutics, to inhibit TERT by competitive interaction with telomeres *via* GQ formation *in vivo*.



T-dGTP

Figure 1.9: *Structure of the deoxyguanosine derivative.T-dGTP.*

Following the publication of details of deoxyguanosine derivatives, which were found to be effective TERT inhibitors, a large volume of research has resulted on deoxyguanosine GQs and their various topologies. To date, the volume of different GQs published is so large that a series of online databases of varying quadruplex structures was established. The goal of this initiative is to combine all known information on GQ formation in the hope of utilising the data in development and applications of GQs in future research.

The Ribosomal Database Project (RDP) was the first such database established by Maidak*et al.* in 1996[45]. In 2006 another online database, utilising a computational approach to map putative quadruplex formation, called the G-rich Sequences Database (GRSDB)[46] was established. The GRSDB was the first database to offer predictive computational software for GQ formation and was designed to study G-quadruplexes near RNA processing sites. Based on the success and high input into GRSDB two new databases were launched in 2007, the GRSDB2 and the G-rich sequence untranslated region database (GRSUTRdb). The GRSDB2 was an upgraded version of its predecessor that incorporated more advanced software for the archiving of data on GQs and observed effects in RNA processing, while the GRS_UTRdb was an entirely new database containing information on the composition and distribution patterns of putative QGs in the 5′- and 3′-UTRs of eukaryotic mRNA sequences[47].

1.8 The growing role of G-quadruplexes

Following the exponential increase in the volume of GQ-related research in the past two decades, it is not surprising that novel applications for GQs have begun to emerge. In recent years GQ formation has been exploited as a method for interaction with ions, providing a delivery system for transport across cell membranes[48]. Further publications have demonstrated the ability of GQs to be used as probes for the qualitative detection of potassium in biological systems. Such research could have a huge impact in the development of early detection approaches for diseases such as multiple sclerosis, which affect the cell`s ability to effectively transmit electrical impulses *via* potassium channels. A wide range of GQs has also been demonstrated to possess peroxidase activity similar to naturally occurring enzymes which protect organisms from harmful oxidation[49]. Nucleic acids are known to play an important role in a wide array of biological process. A key aspect of their function depends upon their ability to fold, and the rate at which they do so. GQs have been used in the development of chip-based biosensors which determine the kinetic folding rate of nucleic acids[50]. GQs have also found highly practical applications, as demonstrated by Li *et al.*, who used GQs in a colorimetric method for the detection of mercury in aqueous solutions[51].

The large body of research on GQs has confirmed that they play key roles in many biological systems, such as the elongation of telomeres. However, the function of GQs is currently not fully understood. Accordingly a large area of research is still based in the detection of GQs. As such, assay kits have been developed that use quadruplex-forming and non-quadruplex-forming oligomer substrates as a method for testing a target compound's ability to form GQs[52]. A notable publication in this regard is the development of an engineered antibody which functions as a probe for the presence of GQs *in vitro*. After the probe antibody binds to the target G-tract sequence *via* GQ formation in human breast adenocarcinoma cells, deep sequencing is utilised to detect and map G-quadruplexes at high resolution. This novel technique has been used to confirm the presence of GQs in sub-telomeres, gene bodies and gene regulatory regions[53].

1.9 Research Motivation

The collective effect of noncovalent factors involved in natural supramolecular systems inspired us to design novel, small molecules based on 2'-deoxyguanosine (*d*G) in order to induce the formation of G-quadruplexes with increased stability and fidelity. It was hypothesized that increasing the number of non-covalent interactions holding the units in a G-tetrad together will stabilize its secondary structure.

The collective contribution toward the over all stability of GQ through a variety of noncovalent interactions such as π-π stacking, hydrophobic forces, Van der Waals forces and electrostatic effects has potential to be further enhanced by expanding the Hoogsteen edge of 2‘-deoxyguanosine. Substitution at C^8 by aromatic groups will lead to a relatively novel class of functionalized nucleosides, 8-Aryl-2'-deoxyguanosine analogues (8ArG) whose supramolecular properties are fine-tuned by the presence and nature of a wide variety of functional groups figure 1.10.

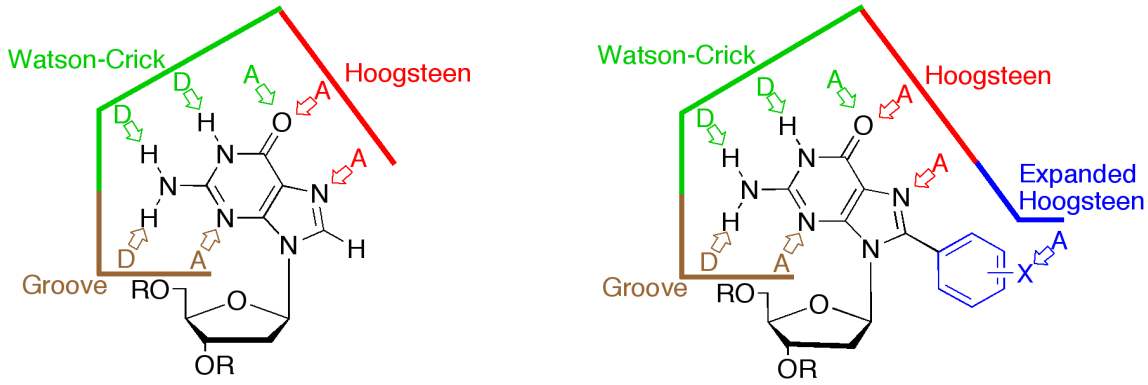


Figure 1.10: *2`-deoxyguanosine molecules self-assemble into GQ using hydrogen bond donors (D) and acceptors (A). On the leftis unmodified 2`-deoxyguanosine which can hydrogen bond on the groove side, Watson-Crick side and the* ***Hoogsteen side.*** *On the right is the 8-aryl-2'-deoxyguanosine derivative and by expanding the Hoogsteen side of 2'-deoxyguanosine with aryl functional groups, we can further stabilise GQ of 2'-deoxyguanosine with addtion hydrogen bond donnors and acceptors.*

In particular, this will be achieved by increasing the number of hydrogen bonds within a G-tetrad by properly placed H-bonding accepting groups on the aromatic ring. Such functional groups in the *meta* position could form up to four more hydrogen bonds with the hydrogen located alongside the "groove edge" of a G-tetrad Figure 1. 11.

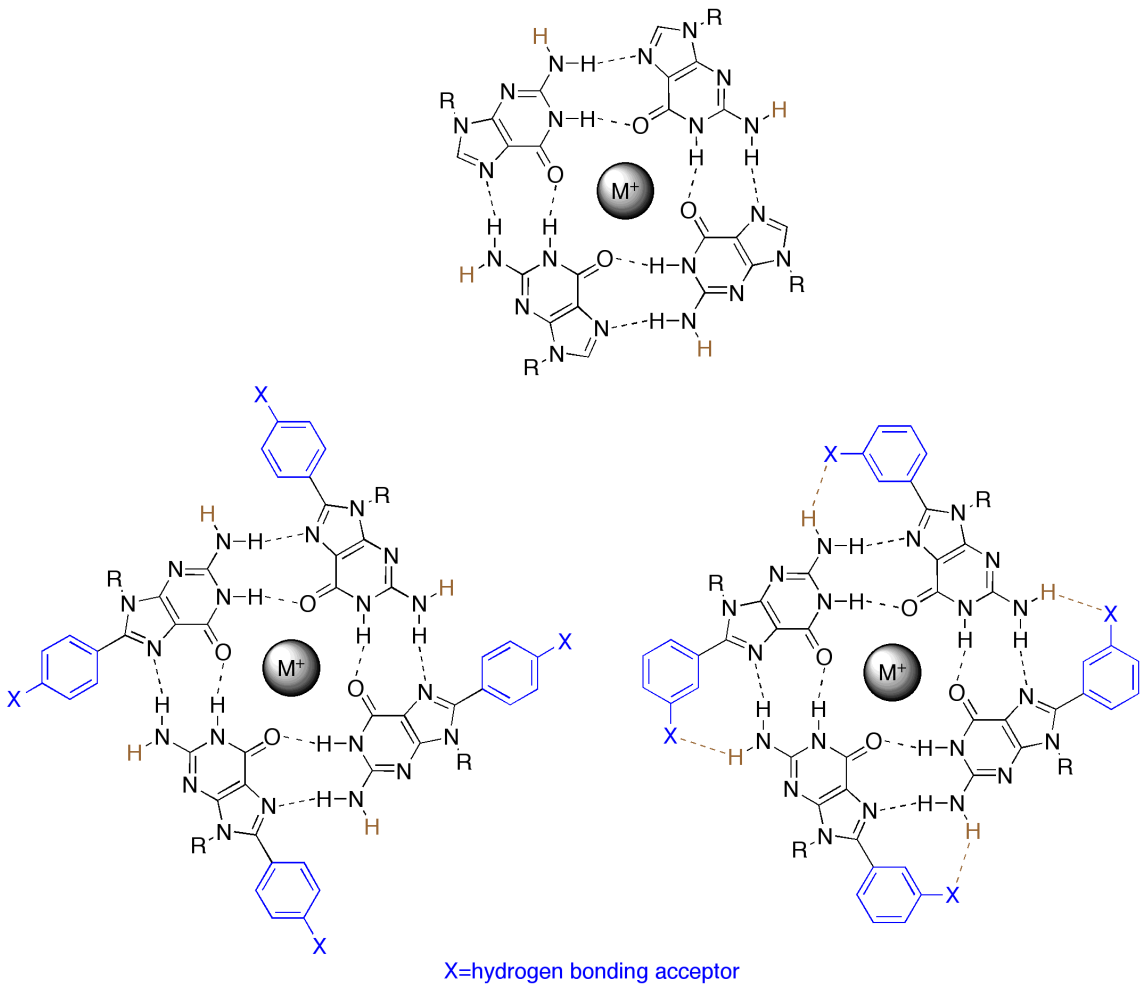


Figure 1.11: *Comparison between the G-tetrads formed by the parent dG (top) with no substitution at position 8 and those formed by 8PhG derivatives. Functional groups in the para position are unable to make hydrogen bonds with hydrogen in the groove (bottom left); functional groups in the meta position can make up to four more H-bonds and thus increase the overall stability of a G-tetrad (bottom right).*

1.9.1 Thesis proposal

The increasing number of applications for GQs in recent years has highlighted the need for a cost-effective method for the rapid screening of the ability of selected small molecules to induce formation and stabilisation of a G-quadruplex. The overall objective of this research therefore is to synthesize a small library of water soluble, 2'-deoxyguanosine derivativesand to develop a novel screening methodology for the ability of small molecules to form and stabilise G-quadruplexes. These objectives will be addressed through the following specific approaches:

1. *Synthesis and full characterisation of a small library of novel compounds based on 2'-deoxygunosine, that are preferably water soluble, with the potential to target G-quadruplex DNA.*

2. *Synthesis of a novel covalently linked G-tetrad. Our hypothesis is that a macrocycle made up of four, covalently cross-linked derivatives of 2'-deoxyguanosine forming one G-tetrad would selectively bind to G-quadruplex with high affinity as it represents one building block of the whole supramolecular structure.*

3. *Development of an optical sensor-based on the quenching phenomenon of folded, labeled DNA on a gold surface. This new technique would allow for rapid screening of large libraries of potential therapeutics and for fast detection of the GQ DNA drug binding events. This technique would characterize those interfaces in a very direct way that delivers the critical information for understanding the basic science and for making a rational choice in modeling and design of new potent chemotherapeutics.*

In the following sections, the rationale and the operational mode of the proposed approach will be explained.

Chapter 2

Materials and Methods

2.1 Materials

Unless otherwise stated, starting materials were obtained in the highest commercial grades and used without further purification.

Purchased from Sigma Aldrich,Vale Road, Arklow, Co Wicklow, Ireland.

4-Dimethylaminopyridine (DMAP), **CAS number: 1122-58-3**

4-Toluenesulfonyl chloride, **CAS number: 98-59-9**

Acetic anhydride, **CAS number: 108-24-7**

Bisacetoxyiodobenzene (BAIB), **CAS number: 3240-34-4**

Carbon tetra bromide, **CAS number:558-13-4**

Chloroform, **CAS number: 67-66-3**

Cyclohexanone-18-crown-6, **CAS number: 17455-13-9**

dimethyl sulfoxide-d_6, **CAS number: 2206-27-1**

Diphenylphosphoryl azide (DPPA), **CAS number: 26386-88-9**

Ethyl acetate (EtOAc), **CAS number: 141-78-6**

Methanesulfonyl chloride, **CAS number: 124-63-0**

Methanol (MeOH), **CAS number: 67-56-1**

Palladium(II) acetate, **CAS number: 3375-31-3**

Pyridine, **CAS number: 110-86-1**

Sodium azide (NaN_3), **CAS number: 26628-22-8**

Sodium bicarbonate ($NaHCO_3$), **CAS number: 144-55-8**

Sodium carbonate, **CAS number: 497-19-8**

Sodium periodate, **CAS number: 7790-28-5**

Triphenylphosphine, **CAS number: 603-35-0**

Goat anti-Human IgG (Fab-specific) **Catalogue number: 12136**

N-bromosuccinimide (NBS), **CAS number: 128-08-5**

Dimethylformamide (DMF), **CAS number: 68-12-2**

Acetonitrile (MeCN), **CAS number: 75-05-8**

Dichloromethane (DCM), **CAS number: 75-09-2**

(2,2,6,6-Tetramethylpiperidin-1-yl)oxy (TEMPO), **CAS number: 2564-83-2**

1,4-Dioxane, **CAS number: 123-91-1**

1,8-Diazabicyclo-undec-7-ene (DBU), **CAS number: 6674-22-2**

3,3′,3″-Phosphanetriyltris(benzenesulfonic acid) trisodium salt (TPPTS), **CAS number 63995-70-0**

Boronic acids purchased from Inochem,Carnforth, Lancashire LA6 1DE., United Kingdom.

1-H-Indazole-4-boronic acid, **CAS number: 1023595-17-6**

1-H-Pyrazole-4-boronic acid, **CAS number: 763120-58-7**

1-H-Pyrazole-5-boronic acid, **CAS number: 376584-63-3**

2-Hydroxyphenylboronic acid, **CAS number: 89466-08-0**

3-Hydroxyphenylboronic acid, **CAS number: 87199-18-6**

4-Carboxyphenylboronic acid, **CAS number: 14047-29-1**

4-ethynylphenyl-boronic acid, **CAS number: 263368-72-5**

4-Hydroxyphenylboronic acid, **CAS number: 71597-85-8**

Uracil-5-boronic acid, **CAS number: 70523-22-7**

Purchased from TCI Europe, Boerenveldseweg 6, Haven 106 B-2070, Zwijndrecht Belgium.

2-Deoxyguanosine, **CAS number: 961-07-9**

Purchased from Eurofins MWG Operon, Anzinger Str. 7a., 85560 Ebersberg, Germany.

30-mer, GC content 50%; 22-mer, GC content 45.5%; and 18-mer GC content 43%, telomeric DNA labelled with cyanine 5 at the 5` position and a thiol group in 3` position, the DNA was purified by HPLC and was comprised of base pair oligonucleotides of sequence TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG.

15-mer, GC content 53.3%, non-telomeric DNA labelled with cyanine 5 at the 5` position and a thiol group in 3` position was purified by HPLC and was comprised of base pair oligonucleotides of sequence ACGGCAGTGTTTAGC.

Purchased from Phasis Ltd, Phasis Sàrl Chemin des Aulx 18, CH – 1228, Plan-les-Ouates Geneva, Switzerland.

30 nm monolayer gold slide, deposited on borosilicate glass.

2.2 Equipment

NMR: ^{1}H and ^{13}C NMR spectra were recorded on Bruker Avance UltraSheild 400MHz spectrometer, with nominal frequencies of 400 MHz for proton or 100 MHz for carbon, respectively. ^{1}H NMR chemical shifts are reported with respect to tetramethylsilane as an internal reference. ^{13}C NMR chemical shifts are reported in parts per million relative to the residual undeuterated dimethyl sulfoxide solvent as an internal reference. The following abbreviations are used to explain the multiplicities: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet and b, broad.

Infrared spectrophotometer: IR was measured on a Perkin Elmer spectrum 100 ATR FT-IR spectrometer. Each sample was measured as solid on a diamond tip plate from 600 to 4000 cm^{-1}, with a scan rate of 1 cm^{-1} per scan.

X-Ray Crystallography: Single Crystal X-ray structures were analysed on an Oxford Diffraction Gemini-S Ultra Diffractometer with Mo radiation at room temperature.

Thin layer chromatography: The TLC analysis was performed on a glass backed Merck Millipore silica gel 60 plates with a pore diameter of 60 Å and F_{254} fluorescent indicator. The mobile phase was a mixture of chloroform 85% (v/v) and methanol 15% (v/v) or dichloromethane 85% (v/v) and methanol 15% (v/v). The samples were measured against the starting materials as standards.

Fluorescent scanner: The fluorescence signal of DNA depositions on gold and Zeonor slides was recorded on a PerkinElmer ScanArray Express (PerkinElmer, MA) with laser excitation wavelength of 633 nm and emission filter wavelength of 670 nm.The fluorescence images were analyzed using Image J software which displays fluorescence emission in the colours white, red green and blue, displaying fluorescence intensity from high emission to low emission respectively.

Dual Polarization Interferometry: The interactions between surface immobilised telomeric DNA and the synthesised 8-Aryl-2'-deoxyguanosine derivatives were measured on a Farfield Analight interferometer 4D workstation (Farfield Scientific Ltd., U.K.). Both thiol-coated and amine-coated Farfield Analight Anachips were used to attach the telomeric DNA via a covalent sulphur-sulphur bond, at concentrations of 1-100 μM.

DNA depositions: The thiol-modified oligonucleotides were deposited on Au-coated (50 nm) slides using a sciFLEXARRAYER S3witha 90 micron Piezo Dispense Capillary (PDC), that deposited droplets of 500 pL volume of thiolated, cyanine 5-labelled, 30-mer telomeric DNA in H_2O in an array of 10, 20 or 50 droplets, respectively.

Mass spectrometer: A Waters Synapt G2 TOF mass spectrometer (Waters, UK) with an electrospray ionisation probe in electrospray positive, acquired data over a mass range 50-800ppm. A lock spray correction was applied to each acquired data set using Leu-Enk.
Inlet conditions, an Acquity UPLC system was coupled at a flow rate of 0.2 ml/min with 50:50 A: Water + 0.1% (v/v) formic acid and B: Acetonitrile + 0.1% (v/v) formic acid so that no chromatography was conducted. Processing was carried out using MassLynx 4.1 constraining the possible elements to show the PPM error for the proposed molecular formula. Samples were made up to approximately 0.25mg/ml in a 50:50 acetonitrile: water + 0.1% (v/v) formic acid.

2.3 Synthetic methods

All reactions requiring anhydrous conditions were carried out using flame-dried glassware under Argon. Reaction progress was analysed using TLC. The isolated products were characterised by NMR, IR spectroscopy, melting point and mass spectrometry.

Synthetic method for 8-bromo-2`-deoxyguanosine (8BrdG) ***Reaction 1***

To a suspension of 2'-deoxyguanosine (4 g, 14 mmol) in acetonitrile (160 mL) and H_2O (40 mL), *N*-bromosuccinimide (NBS, 4 g, 22 mmol) was added in 1 g portions over 15 minutes. The reaction mixture was stirred at room temperature for 40 minutes. The reaction mixture was then vacuum filtered and the precipitate collected. The precipitate was added to acetone (80 mL), sonicated and stirred at room temperature for 2.5 hours. The mixture was then left in a freezer at -2°C degrees Celsius. The mixture was then filtered and washed with acetone (20 mL). The resulting solid was collected and frieze dried to yield 4.36 g (89% yield) of a white solid.

8-Bromo-2'-deoxyguanosine, (8BrdG)

^{1}H-NMR (400 MHz, d6-DMSO) δ 10.79 (s, 1H), 6.48 (s, 2H), 6.17 (dd, J = 6.9 Hz, 1H), 5.22 (br s, 1H), 5.03 (br s, 1H), 4.41 (t, J = 2.9 Hz, 1H), 3.82 (dd, J1 = 8.0 Hz, J2 = 5.3 Hz, 1H), 3.64 (dd, J1 = 11.6 Hz, J2 = 5.3 Hz, 1H), 3.51 (dd, J1 = 11.6 Hz, J2 = 5.3 Hz, 1H), 3.17 (dt, J1 = 13.2 Hz, J2 = 6.9 Hz, 1H), 2.15-2.10 (m, 1H).

^{13}C-NMR (100 MHz, d6-DMSO) δ 155.3, 153.2, 151.9, 120.4, 117.4, 87.8, 85.0, 70.9, 61.9, 36.4.

Synthetic method for Arylation of 8-BrdG ***Reaction 2***

Example: 8-(3-acetylphenyl)-2'-deoxyguanosine

Palladium acetate (9.7 mg, 0.043 mmol), TPPTS (57 mg, 0.101 mmol), sodium carbonate (106 mg, 0.275 mmol), 8-BrdG (500 mg, 1.44 mmol), and the corresponding 3-acetylphenyl boronic acid (2.02 mmol) were placed in a vial under argon. A 25 mL solution of water: acetonitrile (2:1) was degassed and was added to the reaction vessel, the reaction mixture was then purged for 1 min with argon and heated in oil bath at 80 °C until TLC analysis 15% (v/v) methanol: 85% (v/v) chloroform, Rf = 0.3 using starting materials as reference) showed complete conversion. The reaction mixture was then poured into 25 mL H_2O and the pH was adjusted to between 6 and 7 using the drop wise addition of 10% (v/v) HCl. The precipitate formed was stirred at room temp for another 0.5 hour, vacuum filtrated and dried to yield 0.430 g of 8-(3-acetylphenyl)-2'-deoxyguanosine (86% yield) as a white powder.

Note: Synthesis of two additional compounds 2-hydroxy-dG and 4-ethynelphenyl-dG was attempted however, the compounds were not produced in sufficient purity for further analyses.

8-(3-Acetylphenyl)-2'-deoxyguanosine, (3-Ac-dG)

^{1}H-NMR (400 MHz, DMSO-d_6) δ 10.79 (s, 1H), 8.19 (s, 1H), 8.06 (d, J= 7.8 Hz, 1H), 7.90 (d, J=7.8 Hz, 1H), 7.68 (t, J=7.7 Hz, 1H), 6.45 (s, 2H), 6.04 (t, J=7.5 Hz, 1H), 5.13 (d, J=4.3 Hz, 1H), 4.93 (t, J=5.8 Hz, 1H), 4.30 (d, J=2.7 Hz, 1H), 3.77 (q, J=2.9, 5.3 Hz, 1H), 3.61 (m, 1H), 3.52 (m, 1H), 3.32 (s, 3H), 3.17 (m, 1H), 2.02 (ddd, J= 2.7, 6.5, 8.2 Hz, 1H).

^{13}C-NMR (100 MHz, DMSO-d_6) δ 197.97, 157.14, 153.60, 152.57, 146.68, 143.58, 137.51, 133.89, 131.25, 129.57, 123.62, 117.71, 88.35, 85.02, 71.58, 62.51, 39.49, 27.31.

Melting point range: 182 ˚C -186 ˚C

Mass spec: for C18H19N5O5: 386.1464

8-(4-pyrazole-)-2'-deoxyguanosine, (4-pyrazole-dG)

(0.3032 g; 62% yield), 'off' white powder

^{1}H-NMR (400 MHz, DMSO-d_6) 2.102 (m, 1H,), 3.137 (m, 1H,), 3.604 (m, 2H,), 3.819 (s, 1H), 4.440 (s, 1H), 5.190 (m, 2H), 6.380 (s, 2H), 6.761 (s, 1H), 7.257 (s, 1H), 7.897 (s, 1H), 10.862 (s, 1H) 13.255 (s, 1H)

^{13}C-NMR (100 MHz, DMSO-d_6) δ 37.327, 62.376, 71.510, 85.010, 87.834, 105.861, 117.271, 129.628, 140.959, 143.107, 151.757, 152.899, 156.646,

Melting point range: 136 ˚C -140˚C

Mass spec: for C13H15N7O4: 334.1264

8-(5-pyrazole-)-2'-deoxyguanosine (5-pyrazole-dG)

(0.297 g; 61% yield), white powder

^{1}H-NMR (400 MHz, DMSO-d_6) 2.069 (m, H1, H2`), 3.221 (m, 1H, H2`), 3.583 (m, 2H, H5`), 3.811 (t, J=4 Hz, H4`), 4,378 (s, 1H, H3`), 5.057 (s, 1H, 5`OH), 5.236 (s, 1H, 3`OH), 6.159 (t, J=7.2 Hz, 1H, H1`), 6.409 (s, 2H, NH_2), 7.939 (s, 1H, 5H``), 8.078 (s, 1H, 3H``)

^{13}C-NMR (100 MHz, DMSO-d_6) δ 36.408, 61.913, 70.994, 84.152, 87.643, 110.935, 116.764, 141.355, 151.591, 152.739, 156.527,

Melting point range: 132 ˚C -140˚C

Mass spec: for C13H15N7O4: 334.1264

8-(3-hydroxyphenyl) -2`-deoxyguanosine (3-hydroxy-dG)

(0.218 g; 42% yield), white powder

^{1}H-NMR (400 MHz, DMSO-d_6) 2.032 (m, H1, H2`), 3.209 (m, 1H, H2`), 3.627 (m, 2H, H5`), 3.811 (s, 1H, H4`), 4.370 (s, 1H, H3`) 5.044 (s, 1H, 5`OH), 5.201 (s, 1H, 3`OH), 6.106 (t J=7.544 Hz, H1`), 6.433 (s, 2H, NH_2), 6.926 (d, J=7.00 Hz, 1H, 6H``), 7.032 (s, 1H, 2H``), 7.051, (s, 1H, 4H``), 7.341 (t, J=7.8 Hz, 1H, 5H``), 9.826 (s, 1H, NH), 10.820 (s, 1H, OH``)

^{13}C-NMR (100 MHz, DMSO-d_6) δ 36.475, 62.169, 71.288, 84.712, 87.909, 115.867, 116.437, 117.018, 119.731, 129.677, 131.312, 147.146, 151.871, 152.298, 156.298, 157.342,

Melting point range: 212˚C -215˚C

Mass spec: for C16H17N5O5: 360.1308

8-(4-hydroxyphenyl) -2`-deoxyguanosine (4-hydroxy-dG)

(0.412 g; 80% yield), white powder

^{1}H-NMR (400 MHz, DMSO-d_6) 2.011 (m,H1, H2`), 3.160 (m, 1H, H2`) 3.603 (m, 2H, H5`), 3.791 (s, 1H, H4`), 4.340 (s, 1H, H3`), 5.062 (s, 1H, 5`OH), 5.192 (s, 1H, 3`OH), 6.052 (t, J=7.20 Hz, 1H, H1`), 6.399 (s, 2H, NH_2), 6.897 (d, J=8.4 Hz, 2H, H2`` H6``), 7.454 (d, J=8.4 Hz, H3`` H5``), 9.974 (s, 1H, NH), 10.750 (s, 1H, OH``)

13C-NMR (100 MHz, DMSO-d_6) δ 37.606, 60.971, 70.057, 86.623, 114.198, 115.661, 119.718, 129.499, 146.377, 150.619, 151.660, 157.392,

Melting point range: 222˚C -228˚C

Mass spec: for C16H17N5O5: 360.1308

General synthetic method for Esterification of 8-ArdG ***Reaction 3***

8-Ar-dG (0.723 mmol) pre-dried by suspension in acetonitrile and solvent evaporation (3x) was suspended in anhydrous acetonitrile under nitrogen. To this suspension, TEA (153.47 mg, 1.516 mmol), acetic anhydride (154.84 mg, 1.516 mmol) and DMAP (9 mg, 0.072 mmol) were added. The reaction mixture was quenched by adding excess of MeOH followed by solvent evaporation. The resulting solid was dissolved in EtOAc and washed with 10 % (w/v) $NaHCO_3$ (2 x 15 ml) and brine (1 x 15 ml). The organic phase was separated and solvent removed by rotary evaporation. The resulting crude was analysed by thin layer chromatography. However the analyses showed no new products, only the presence of the starting materials.

Preparation of 4-pyrazole-dG crystals

The crystal structure of the 4-pyrazole-dG compound was obtained by dissolving 100 mg of the compound in a 400 ml solution of acetonitrile and water in 1:1 ratio in a 1 litre evaporating dish. The evaporating dish containing the mixture was covered with punctured tinfoil and place in a fumehood for three weeks. The solvent was left to evaporate at room temperature. After three weeks, crystals were observed to have grown.

2.4 Biological studies

18-mer, 22-mer and 30-mer, telomeric DNA were modified with a 3'-thiol group and a 5`-cyanine-5 fluorophoreemitting at 670 nm. All DNA samples were purchased modified to specification from Eurofins MWG Operon. The synthesis report is attached in the appendix. All DNA samples were prepared in a sterile environment in close proximity to an activated Bunsen burner.

2.4.1 Protocol for de-protection of the 3`-thiol terminal group.

To an Eppendorf tube containing the telomeric DNA 30-mer as a dried pellet 200 µL of 10 mM tris(2-carboxyethyl)phosphine(TCEP) was added. The mixture was sonicated for 60min at room temp, 150 µL of 3 M sodium acetate added to the mixture and the tube containing the DNA mixture was filled with 2.5 mL ethanol and shaken gently. The tube was incubated for 20 min at -20°C. The tube was inserted in a centrifuge and spun for 5 min at 7000 rpm on a thermo scientific Heraeus Pico 17 centrifuge (2.739g). The supernatant was discarded and the pellet was air dried at room temp.

2.4.2 General Procedure for telomeric *DNA – 8ArdG derivatives interaction studies (DPI experiments).*

The DPI experiments were performed on Analight Bio 200 amine and thiolated Anachips purchased from Farfield. The mobile phase was a phosphate buffered saline solution (PBS) pH 7.4 and 10 mmol concentration. The instrument was run at a flow rate of 100 µL/min until equilibration of the baseline was established. Then the experiment was initiated. At the beginning of all experiments, the DPI was calibrated to the surface of the Anachip to be used in that experiment, using the refractive index of degassed, high purity ethanol and de-ionised water, filtered through a Millipak 0.22 µm membrane filter as standards. For all experiments the DNA was incubated at 10 µM concentration in PBS. The flow rate for all experiments was set at 100 µL/min unless otherwise stated. All injections were performed with a volume of 440 µL at a flow rate 5 µL/min. Compounds 4-hydroxy-dG, 3-hydroxy-dG, 4-pyrazole-dG, 5-pyrazole-dG were dissolved at a concentration of 100 µM in a mixture of H_2O/MeCN in 1:1 ratio. Version 1.7 of Farfield Analight software was used for the analysis of the crude data. Injection times for DPI experiments were presented in the format of hours:minutes:seconds attached to the DPI display in the appendix section.

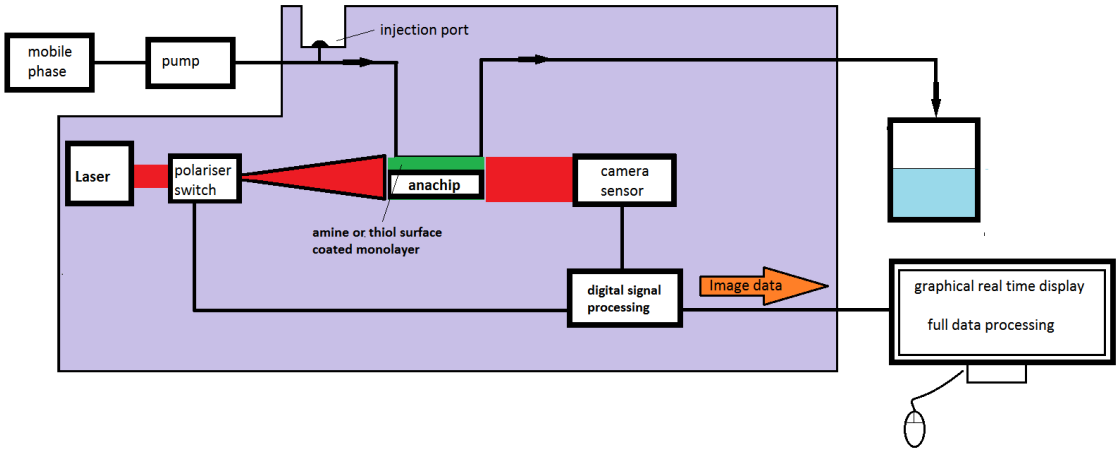


Figure 2.1: *Cartoon illustrating the DPI system. The Anachip is inserted in a mount and locked in place. Samples are injected into the system via the injection port, once injected the sample is carried in the mobile phase to the surface of the Anachip, polarised light is passed through the Anachip to a camera sensor, the camera is connected to digital signal processing unit which sends real time image data of the Anachip surface to a computer.*

Thiol-coated Anachip experiments

After calibration, to allow the DNA to bind to the thiol surface, the DNA was injected at 10 µM concentration and, after a few seconds, the flow rate was set to 0 µL/min for a 10 hour incubation period, followed by 440 µl of 5 mM KI in H_2O. Then, a 440 µl solution of 100 µM KI and 100 µM of 3-hydroxy-dG or 4-hydroxy-dG, in a 1:1 MeCN/H_2O were injected. The resulting changes in thickness on the chip surface were recorded.

Amine-coated Anachip experiments

After calibration, freshly prepared 5 mM sulfo-SMCC linker in H_2O was injected prior to the injection of DNA at 10 µM concentration. Then, a 440 µl solution of 100 µM KI and 100 µM of 4-pyrazole-dG or 5-pyrazole-dG, in a 1:1 MeCN/H_2O was injected. The resulting changes in thickness on the chip surface were recorded.

General procedure for DNA depositions on Zeonor® and gold slides

The DNA depositions were performed using a sciFLEXARRAYER S3 instrument (Figure 2.2) DNA was deposited in ordered arrays on the surface of the slides using a 90 micron Piezo Dispense Capillary (PDC). The total volume of each drop was ~ 500pL. The depositions were performed using a variable number of drops per slide. The number of droplets deposited per slide was 20, 30 or 50. Cyclic olefin polymers (Zeonor® 1060R microscope slides), of dimensions 25mm by 75mm by 1.1mm, were purchased from Sigolis AB. The Zeonor slides were plasma-activated using an Oxford-Instruments PlasmaPro 100 PECVD chamber, with an operating frequency of 13.56 MHz. Telomeric, cyanine 5, 30-mer DNA in H_2O was deposited on to the surface of the Zeonor slide, immediately after the slide surface was oxidised. The

DNA was allowed to incubate on the surface for two hours. After incubation the slides were washed with a 1% (w/v) solution of sodium dodecyl sulphate. Borosilicate glass slides, with a 30nm uniform gold layer of the dimensions 25mm by 75mm by 1.1mm were purchased from Phasis Ltd. The telomeric DNA was deposited on the gold-coated slides according to the same protocol as described above. The DNA-spotted gold slides were thoroughly rinsed and then allowed to incubate in the solution of the synthesised 8ArdG derivatives for two hours. The solutions were then carefully removed with the use of a vacuum line. The slides were allowed to dry, while protected from light to prevent photo bleaching of the dye. The fluorescence images are presented in the results in section 4.2. The fluorescence emission images of the slides were measured using a Perkin Elmer fluorescence scanner at an excitation wavelength of 650 nm and an emission wavelength of 670 nm at 70% gain.

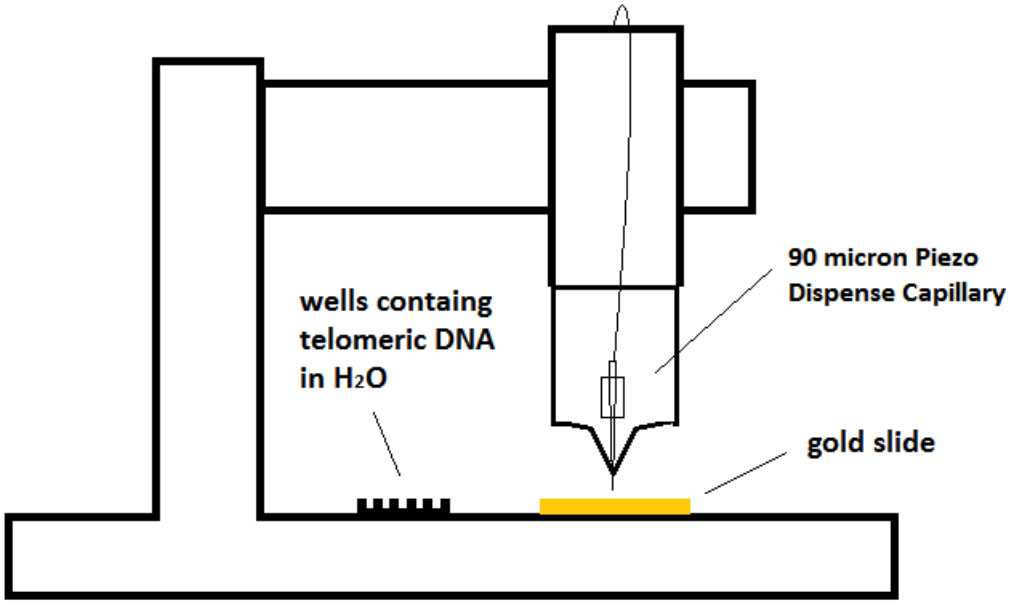


Figure 2.2: *Diagram illustrating a sciFLEXARRAYER S3 instrument depositing droplets of DNA in H_2O on the surface of a gold slide. The 90 micron Piezo Dispense Capillary draws the sample DNA from wells mounted on the platform of the instrument.*

Chapter 3

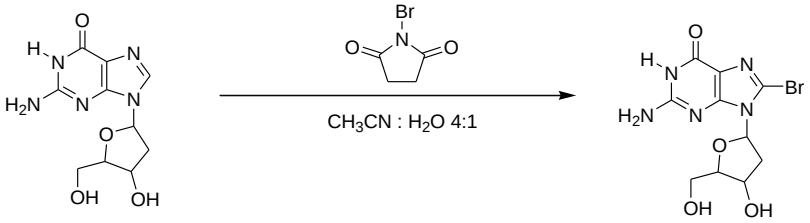
Synthesis

3.1 Results

3.1.1 Synthesis of a 8-Aryl- 2`-Deoxyguanosine Derivatives

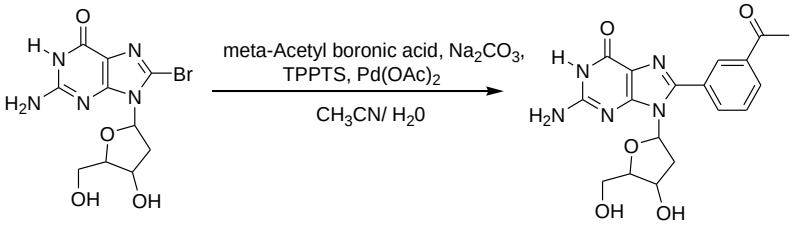
3.1.1.1 Reaction 1: Synthesis of 8-Bromo-2`-deoxyguanosine

The first aim of the project, as outlined in the thesis proposal (section 1.9.1), was to synthesise a library of novel 8-Aryl-2`-deoxyguanosine derivatives. To facilitate the coupling of functional groups, a halide leaving group was incorporated in the precursor molecule. The compound was modified by the addition of a bromide in the 8-positionof deoxyguanosine, which serves as an excellent leaving group. The precursor 8-bromo-2`-deoxyguanosine was prepared by nucleophilic aromatic substitution bromination using *N*-bromosuccinimide as shown in scheme 3.1. The 8-bromo-dG was isolated in 87% yield.

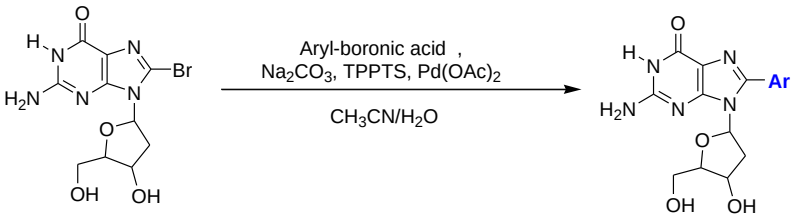


Scheme 3.1: **Synthesis of 8-bromodeoxyguanosine.**

The first synthesised compound was the previously reported 8-(3-acetylphenyl-2`-deoxyguanosine. The self-assembling properties of this compound were previously reported and we aimed to use it as a reference molecule.



Four additional compounds were selected as potential candidates for synthesis to be used in further analysis in the second phase of the project. The target compounds were selected on the basis of examining the varying topologies of GQ assembly in the telomeric DNA. The cost of starting materials and ease of synthesis were also contributing factors in the selection of the target compounds. The purity of all compounds was assessed by ^{1}H Nuclear Magnetic Resonance (NMR) and high resolution mass spectrometry. All target compounds were synthesised according to the general reaction scheme 3.2. The corresponding yields and reaction times are summarized in table 3.1.

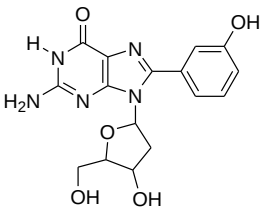


Scheme 3.2: *Proposed library of novel compounds of 2`-deoxyguanosine derivatives.*

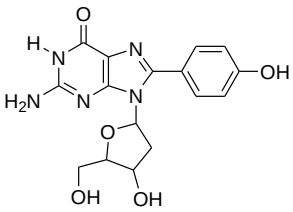
Synthesis of 2-hydroxyphenyl-dG proceeded with the formation of by-product impurities. Purification of 2-hydroxyphenyl-dG was attempted *via* column chromatography, however, the crude product proved highly insoluble and a suitable mobile phase for 2-hydroxyphenyl-dG was not discovered. Removal of impurities by solvent washing was repeatedly attempted. However, the polarities of the impurities were too similar to that of the product and solvent washing proved unsuccessful. Of the five target compounds, four compounds were successfully synthesised in sufficient purity for use in biological assays. (Figure 3.1).

5-pyrazole-dG

4-pyrazole-dG



3-hydroxyphenyl-dG

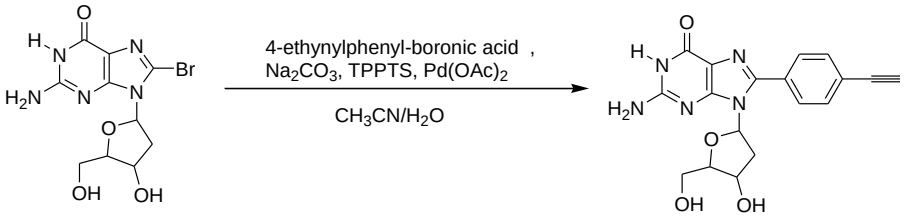


4-hydroxyphenyl-dG

Figure 3.1: *Library of synthesised derivatives of 2`-deoxyguanosine.*

The final library of compounds was 3-hydroxyphenyl-dG, 4-hydroxyphenyl-dG, 4-pyrazole-dG and 5-pyrazole-dG. The synthetic details and the corresponding yields are summarised in table 3.1.

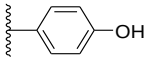
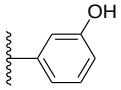
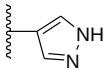
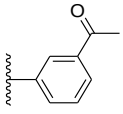
A coupling reaction between 4-ethynyl boroanic acid and 8BrdG was performed in an attempt to synthesis a precursor molecule, intended for the synthesis of a covalently linked G-tetrad Scheme 3.3.



Scheme 3.3: *Proposed synthesis for the arylation of 2`-deoxyguanosine in the C8 position using aryl boronic acids.*

The resulting crude was analysed by thin layer chromatography. However the analyses showed the presence of a great many undesired products. The crude products proved insoluble in acetone, pentane, hexane, ethyl acetate, chloroform, dichloromethane, methanol, ethanol and diethyl ether. Accordingly purification of the crude product was abandoned and the target product was not identified and isolated.

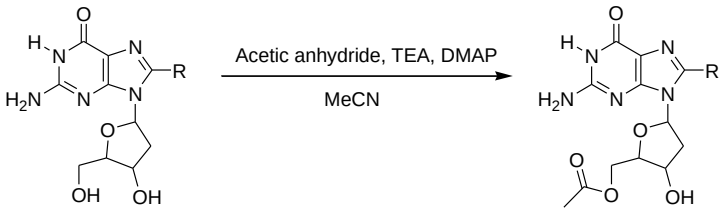
Table 3.1 Products of Suzuki coupling

Name	Reaction type	Time (hrs)	**Aryl** functional group	Isolated yield (%)
4-hydroxy-dG	Suzuki	10		80%
3-hydroxy-dG	Suzuki	10		42%
4-pyrazole-dG	Suzuki	6		62%
5-pyrazole-dG	Suzuki	6		61%
3-Ac-dG	Suzuki	6		86%

3.1.2: Reaction 3: Synthesis of library of compounds of 2`-deoxyguanosine derivatives with esterification of 5`OH

The addition of an aromatic group to the 2`-deoxyguanosine significantly changed the overall electron density of the compounds and enhanced the capacity of further π-π stacking interactions. It was therefore reasoned that the modifications to the 2`-deoxyguanosine could

affect the polarity of the molecules to a sufficient degree to increase solubility in low boiling organic solvents. However, the solubility of the 8ArdG derivatives remained poor in acetone, pentane, hexane, ethyl acetate, chloroform, dichloromethane, methanol and ethanol. In order to increase solubility, further functionalization of the four compounds was attempted, to introduce lipophilic acetyl group on the 5' OH group on the deoxyribose (Figure 3.4). Previous reports by Gubala *et al.*[54] showed that conversion of the 5` -OH group on the deoxyribose to the corresponding isobutyl esters significantly improved solubility in chloroform. To improve the solubility of the library of compounds, the acetylation of the crude products to their respective methyl esters was attempted *via* a catalytic esterification reaction, modelled on the Yamaguchi esterification, as shown in scheme 3.4.



Scheme 3.4: *Proposed esterification reaction of the 5` hydroxyl group on the library of compounds.*

Previous reports demonstrated that it was critical that the reactions must be performed in an anhydrous environment[55]. Accordingly, the reactions were performed in a nitrogen glove box in anhydrous acetonitrile. Unfortunately, the compounds proved to be insoluble and TLC analysis of the reaction mixture did not demonstrate the target acetylated compounds. The list of attempted reactions and the results are summarised in table 3.2 see below. The reactions are further discussed in section 3.4.2.

Table 3.2 Products of esterification reaction

Name	Reaction type	Time (hrs)	Substitution position	Functional group substitution	Isolated yield (%)
4-hydroxy-5`-Ac- dG	Esterification	10	$C^{5`}$		No product
3-hydroxy-5`-Ac- dG	Esterification	10	$C^{5`}$		No product
4-pyrazole-5`-Ac-dG	Esterification	10	$C^{5`}$		No product
5-pyrazole-5`-Ac-dG	Esterification	10	$C^{5`}$		No product

3.2 Covalently Linked G-Tetrad

3.2.1 Proposed synthetic route

The next goal of the project, as stated in the thesis proposal, section (1.9.1), was the preparation of a covalently linked G-tetrad (CLGT) illustrated in Figure 3.3. To date no such covalently linked tetrad has been published. The self-assembly of GQ of deoxyguanosine derivatives is possible due to the presence of two hydrogen bond donors and three hydrogen bond acceptors as shown in Figure 3.2. The goal was to utilise a tailored novel derivative of deoxyguanosine to assemble *via* hydrogen bonding into a GQ in solution, and then perform the covalently linking reaction *in situ*.

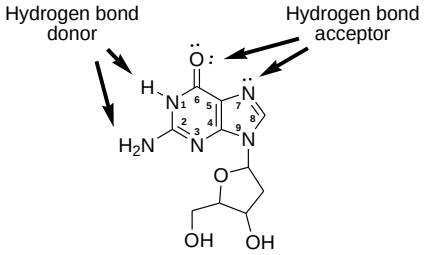


Figure 3.2: *Required sites of 2`-deoxyguanosine for hydrogen bonding.*

The hydrogen bond donors are the secondary amide (numbered as 1 in Figure 3.2) and the primary amine adjacent to carbon 2. The two hydrogen bond acceptors are the tertiary amine (7) and the oxygen of the carbonyl group (6). The 3` and 5` hydroxyl groups on the cyclic furanose sugar do not participate in the formation of 2`-deoxguanosine G-tetrads[56]. Therefore, the carbon in the $C^8$ position and 5` hydroxyl groups, are viable targets for substitution of functional groups that could be allowed to undergo further subsequent reactions to covalently stabilise the G-tetrad (Figure 3.3). The fictionalisations were to be performed after the induced G-tetrad formation in solution *via* the addition of a cation to the novel modified 2`-deoxguanosine molecules.

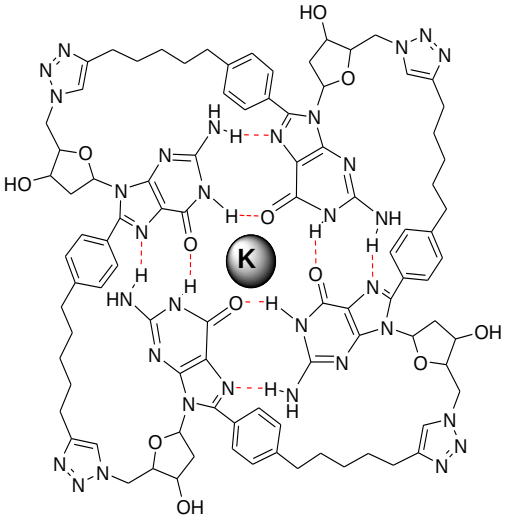
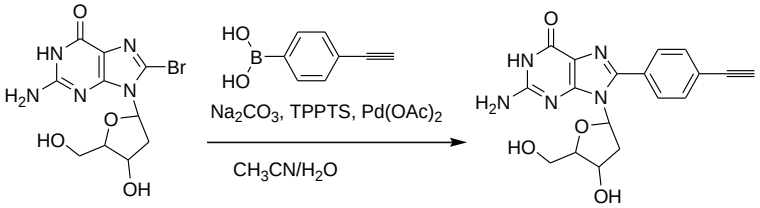


Figure 3.3: *Proposed covalently-linked G-tetrad.*

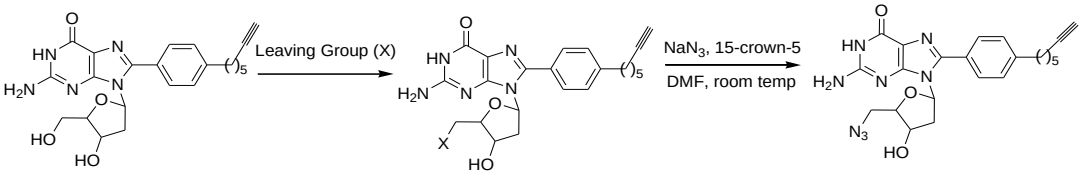
The proposed synthesis of the CLGT was dependant on the synthesis of a novel derivative of 2`-deoxyguanosine, functionalized in the $C^8$ position with an aryl alkyne and in the 5` position with an azide. Before the materials for the proposed synthesis depicted in figure 3.3 were purchased, a trial synthesis was performed with a similar, available boronic acid starting material, 4-ethynylphenyl boronic acid. Synthesis of a 4-ethynylphenyl-dG derivative was attempted for the synthesis of CLGT (Scheme 3.5). The synthesis of the compound proved problematic and a multitude of undesired by-products were observed in the TLC analyses of the crude product.



Scheme 3.5: *Functionalization of C^8 position to aryl-alkyne.*

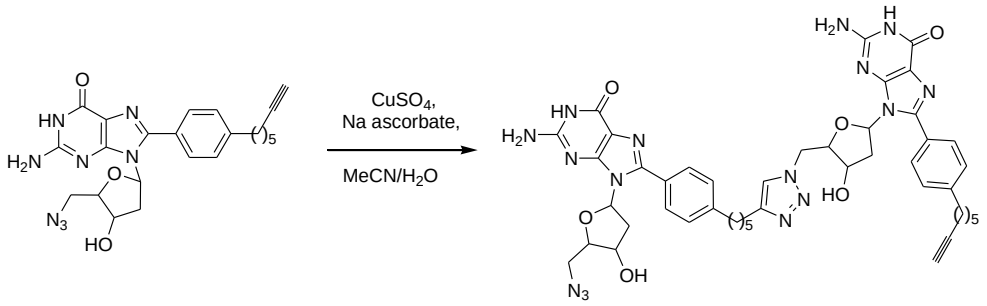
3.2.2 Click chemistry as a method for covalent linking

The chosen synthetic method for the synthesis of the CLGT was based on click chemistry. Click chemistry is a term used to describe a host of chemical reactions that quickly and reliably join small molecules together. The reactions were designed to mimic similar reactions found in nature and are relatively simple to perform. The concept of click chemistry was first fully described by Barry Sharlpless and Valery Folkin in 2001[57]. The proposed synthesis involved an azide-alkyne Huisgen cycloaddition to form the G-tetrad illustrated in Scheme 3.5. The proposed click chemistry reaction was to be performed in two steps. The first step was to substitute the 5` OH with an azide functional group (Scheme 3.6).



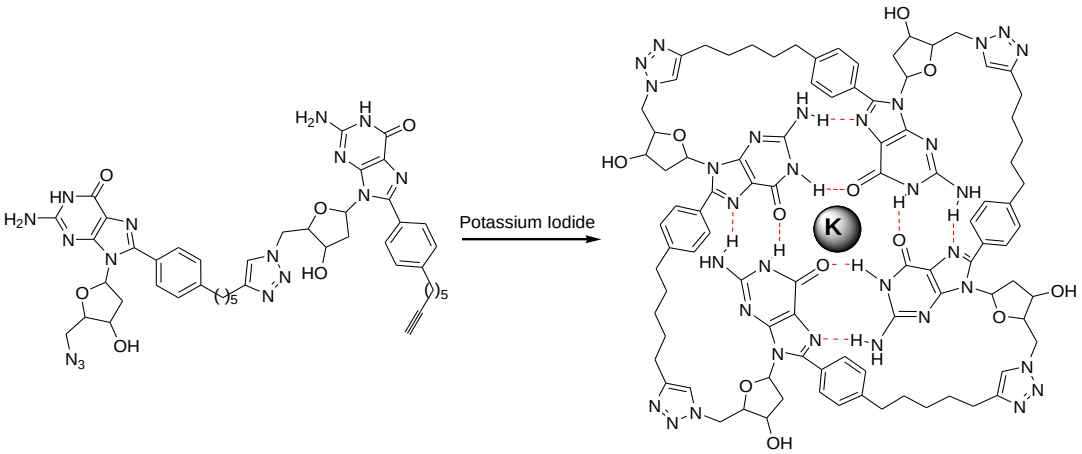
Scheme 3.6 *Proposed preparation of 5`azide with an aryl ethynyl group in position C[8].*

The 5` position of the sugar is relatively unhindered and was reasoned to be further available for subsequent reactions after the substitution of the azide. The furanose sugar is connected to the guanine base *via* a glycosidic bond. Although 2`-deoxyguanosine favours the *cis* conformation, the sugar can rotate around the glycosidic bond further increasing the likelihood that the substituted 5`azide would be left unhindered, thus allowing covalent linkage to the G-tetrad *via* click chemistry. The second step of the proposed click chemistry reaction was the coupling the ethynyl aryl functional group in the $C^8$ position (Scheme 3.7) to the 5` azide, forming a cyclic triazole linker between the two molecules, i.e. an Huisgen cycloaddition[58].



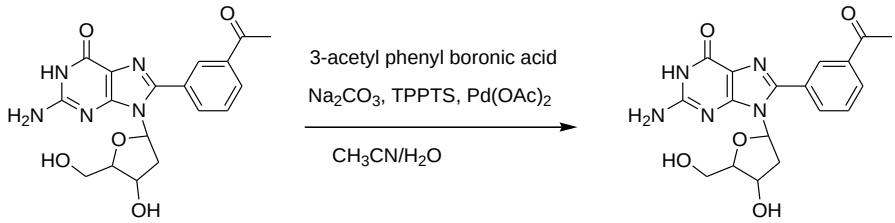
Scheme 3.7: *Formation of covalent link via an azide-alkyne Huisgen cycloaddition.*

The covalent linking of the four subunits into the CLGT via the Huisgen cycloadition of an alkyne to an azide, was to be performed while the novel 2`-deoxyguanosine compounds assembled into a tetrahedral conformation around a metal cation. It was postulated that the coordination of the compounds around the cation would provide a favourable orientation for the click chemistry coupling reaction. Potassium iodide was to be added with the click chemistry reactants. However, for clarity, the coupling reaction and the self-assembly of the CLGT have been presented separately as schemes 3.7 and 3.8.



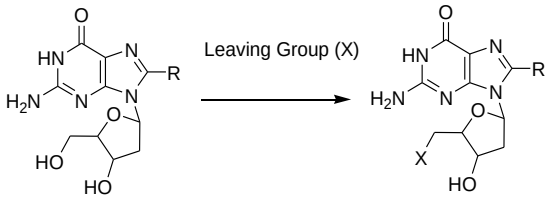
Scheme 3.8: *Covalent linking of the azide-alkyne via an Huisgen cycloaddition.*

The first compound synthesised for production of the CLGT was a previously reported compound called 8-(3-acetylphenyl)-2`-deoxyguanosine (3-Ac-dG), as shown in Scheme 3.8. This compound was synthesised in high yield with purity and, unlike most nucleosides, is soluble in a 1:1 mixture of water and acetonitrile and a 1:1 mixture of methanol and dichloromethane. Due to its favourable solubility, it was used as starting material in all CLGT trial reactions.



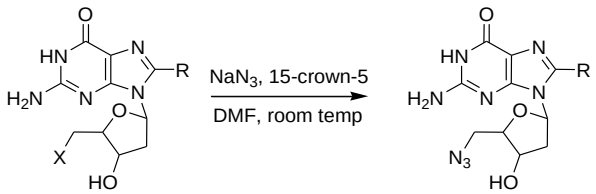
Scheme 3.9: *Synthesis of 3-Ac-dG via Suzuki coupling.*

As previously mentioned, click chemistry coupling reactions are known for their reliability. Thus, the most challenging aspect of the synthesis was the successful substitution of the 5`OH for the 5`azide. The functionalization of the 5` OH group to an azide was intended to be performed as a two-step reaction. The first reaction was intended to synthesise an intermediate compound that would replace the OH group for a good leaving group, X (Scheme 3.10).



Scheme 3.10: *Conversion of 5'OH to a good leaving group (X).*

The second reaction was intended to be performed on the intermediate compound, an S_N2 substitution of the leaving group for an azide, as displayed in Scheme 3.11.



Scheme 3.11: *Substitution reaction of leaving group X for azide.*

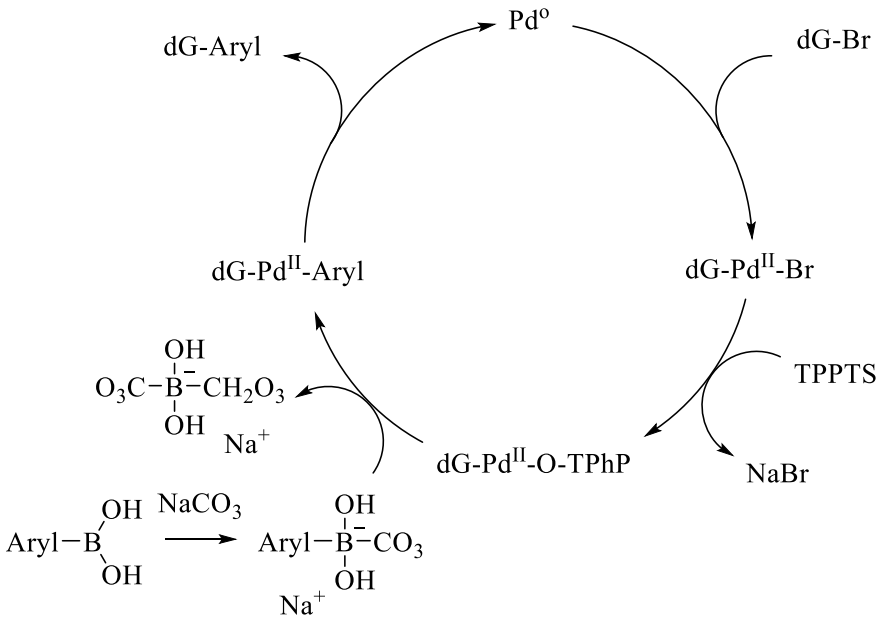
In an attempt to synthesise the intermediate compound, it was essential to find a leaving group that would successfully replace the 5`OH. A wide variety of substituents that typically serve as excellent leaving groups for S_N2 substitutions were considered. The reactions performed are described in the following sub-sections.

3.3 Discussion

In selecting an appropriate coupling reaction for the synthesis of the library of compounds, the following list of criteria must be satisfied. Firstly, unprotected nucleic acids such as deoxyguanosine possess extremely poor solubility in most organic solvents such a methanol, ethanol and dichloromethane, therefore, it was essential the reactions could proceed in a solvent from which they could be recovered. Second, the reaction needed to be efficient at generating a carbon-carbon bond between 2`-deoxyguanosine and prospective substrates. Third, it was essential the reaction could proceed in an aqueous mixture, as the work up used in the synthesis of the desired target products employed precipitation of the 2`-deoxyguanosine derivatives from the crude reaction mixture which remained solvated in the aqueous mixture. A reaction that meets all three of these conditions was the Suzuki coupling reaction. The Suzuki coupling reaction was selected based upon on the successful coupling of 2`-deoxyguanosine derivatives reported within the literature[60]. Additionally, the catalyst for this reaction, palladium, had water soluble derivatives commercially available. Most important of all, however, was that there exists a wide variety of functionalized boronic acids commercially available, which would provide the functional groups to be coupled to the nucleoside.

3.3.1 Suzuki Coupling Reaction

In 1979 Akira Suzuki *et al.* demonstrated a new stereospecific cross-coupling by the palladium-catalyzed reaction of alkenylboranes with 1-alkenyl and 1-alkynyl halides[61,62]. This was the birth of the Suzuki coupling reaction, a milestone in organic synthesis, as it provided a robust and versatile methodology for coupling reactions (the formation of carbon to carbon bonds) between complex aryl systems, a feat which, prior to the Suzuki reaction, was extremely difficult to achieve. The reaction is highly sensitive to oxygen, therefore the reaction must be performed under oxygen free conditions. The reason for this is that the palladium (II) acetate catalyst is highly sensitive to oxidation. If the reaction is exposed to air, the palladium will oxidise to the palladium (0) oxidation state. In this state the palladium cannot facilitate the transmetalation step required to transfer the boronic acid aryl ligand to the target nucleotide. The Suzuki coupling reactions were performed under nitrogen. The Suzuki reaction is a coupling reaction where the coupling partners are a boronic acid with a halide catalysed by a palladium (0) complex. The catalytic scheme of the Suzuki coupling reaction and a discussion of its mechanism are displayed below in Scheme 3.21.



Scheme 3.21: *Catalytic cycle of Suzuki coupling. Pd^0 is the palladium acetate catalyst in its (0) oxidation state. dG-Br is the 8-bromo-2`-deoxyguanosine, halide species. TPPTS is the 3,3',3''-Phosphanetriyltris (benzenesulfonic acid) trisodium salt responsible for the transfer of ligands between boronic acid and the 2`-deoxyguanosine species. While the halide is coupled to the palladium catalyst, it exists in the 2^+ oxidation state and is displayed as dG-Pd^{II}-Br. The halide is substituted for triphenyl phosphine ligand after a transmetalation step, displayed as dG-Pd^{II}-O-TPhP. dG-Pd^{II}-aryl represents the Aryl group coupled to the palladium in 2^+ oxidation state. dG-Aryl is the target Suzuki coupled product of the aryl boronic acid and the substrate starting material.*

The mechanism of the Suzuki reaction begins with the addition of the palladium catalyst. The first step is the oxidative addition of palladium. The halide attacks palladium followed by the nucleoside to form a palladium $(II)^+$ species. The oxidative addition to palladium will proceed with retention of stereochemistry for vinyl halides but will produce inversion of stereochemistry with allylic and benzylic halides[63.] The oxidative addition initially forms the *cis*–palladium complex, which rapidly isomerizes to the trans-complex[64]. The reaction of the organohalide-palladium (II) species with the 3,3′,3″-phosphanetriyltris(benzenesulfonic acid) trisodium salt (TPPTS) substitutes the halide for the phosphine ligand. The sodium carbonate base attacks the boronic acid to form the boronate salt complex which reacts with the organopalladium species in a transmetalation reaction. Transmetalation is an organometallic reaction that involves the transfer of ligands from one species to another. In Suzuki couplings the aryl ligands are transferred from the boronic acid to acetyl palladium (II) species. Reductive elimination of the desired product restores the palladium to the original palladium (0) catalyst. Suzuki coupling has a broad scope in relation to the various arylated substituents that can be substituted to the nucleoside. Another advantages is that it can facilitate coupling reactions in a mixture of organic

and aqueous conditions which was essential for this project given that primary starting material 2`-deoxyguanosine is insoluble in pure H_2O and in the majority of low boiling organic solvents. In 2010 Akira Suzuki was jointly awarded the Noble prize with Richard F. Heck and Ei-ichi Negishi in chemistry for his work in 1979 on palladium cross coupling reactions. The products of the Suzuki-mayora were characterised by mass spectrometry, IR, NMR and melting point.

Infra Red Spectroscopy

For the compounds 3-hydroxy-dG, Figure ***A1.1*** and 4-hydroxy-dG Figure ***A1.2*** the OH band at 3163cm^{-1} is quite large . The peak is generally braod due to the presence of the 3` and 5` OH. However, the additional meta and para OH on the phenyl group of 3-hydroxy-dG and 4-hydroxy-dG, respectively, results in a characteristic deeper broader peak, as can be observed in the spectra included in the appendix section 2. Both spectra also show the characteristic peaks for the presence of NH groups at 3406 cm^{-1}, sharp peaks at 1630 cm^{-1} that indicate the presence of nitrogen linked carbonaly groups and sharp peaks at 1594 cm^{-1} that are indicative of aromatic groups.

Analyses of the IR spectra of compounds 4-pyrazole-dG Figure ***A1.3*** and 5-pyrazole-dG Figure ***A1.4*** show a sharp peak at 3406 cm^{-1} that correspond the presence of additional NH groups. Both of the spectra also display sharp peaks at 1630 cm^{-1} that indicate the presence of nitrogen linked carbonyl groups, sharp peaks at 1594 cm^{-1} indicative of aromatic groups and broad peaks at 3163cm^{-1} confirming the presence of the 3` and 5` OH groups.

Nuclear magnetic resonance spectroscopy

The compounds 3-hydroxyphenyl-dG, 4-hydroxyphenyl-dG, 4-pyrazole-dG, 5-pyrazole-dG and 8-(3-Acetylphenyl)-2'-deoxyguanosine were fully characterised by ***^{1}H, ^{13}C, HMBC and HMQC*** nuclear magnetic resonance spectroscopy (NMR). The characteristics of NMR spectra confirmed the identity of the respective compounds.

Mass spectrometry

The mass fargements produced for the compounds 3-hydroxyphenyl-dG, 4-hydroxyphenyl-dG, 4-pyrazole-dG, 5-pyrazole-dG and 8-(3-Acetylphenyl)-2'-dG under mass spectrum analyses correspond to the respective theoretical calculated values.

3.3.2 Synthesis of 8-(2-Hydroxyphenyl)-dG

The synthesis of the library of compounds was initially intended to include 3 isomers of hydroxyphenyl-dG series. However, the synthesis of 2-hydroxyphenyl-dG proved problematic. The target compound was synthesised using the procedure described in section 2.3. The ^{1}H NMR revealed the presence of an unacceptable quantity of by-products, see appendix, Section 2, figure A2.22. In contrast, the synthesis of compounds 8-(3-hydroxyphenyl)-dG and 8-(4-

hydroxyphenyl)-dG were performed successfully with no impurities. Our hypothesis for this result is that perhaps the steric orientation of 8-(2-hydroxyphenyl)-dG may be unfavourable. The successful coupling of the 2-hydroxyphenyl boronic acid to the 8-bromo-dG would require the hydroxyl group of the boronic acid to come in close proximity to the amide in position 7 of the 8-bromo-dG (Figure 3.5). The close proximity of the lone pair of electrons on the nitrogen and oxygen could have caused an unfavourable increase in electron density and repulsions between the two molecules from one another, resulting in the increased formation of undesired by-products.

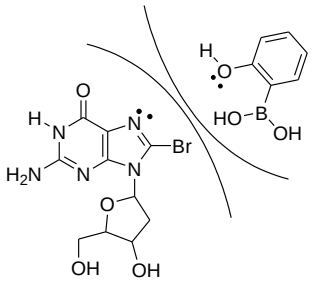


Figure 3.5: *Steric repulsion of lone pairs from position 7 amide of 8-bromo-deoxyguanosine and position 2 hydroxyl of 2-hydroxyphenyl boronic acid group.*

3.3.3 Unusual ^{1}H NMR characteristics of 4-pyrazole-dG

The 4-pyrazole-dG analogue displayed surprising characteristics in two dimensional NMR analysis. This novel compound has displayed unusual chemical shifts attributed to a previously unreported molecular conformation (Figure 3.6). 2`-deoxyguanosine and derivatives of 2`-deoxyguanosine, typically display 1H` proton peak 1ppm upfield from the $NH_2$ peak outlined in **blue**. The 1H` proton on the furanose sugar and the 5``NH proton on the pyrazole ring have been observed to be interacting with one another. The evidence for this interaction appears in the form of significant de-shielding of the 1H` proton outlined in **red**, resulting in the characteristic peak of the 1H` proton appearing over 2ppm downfield of its expected location.

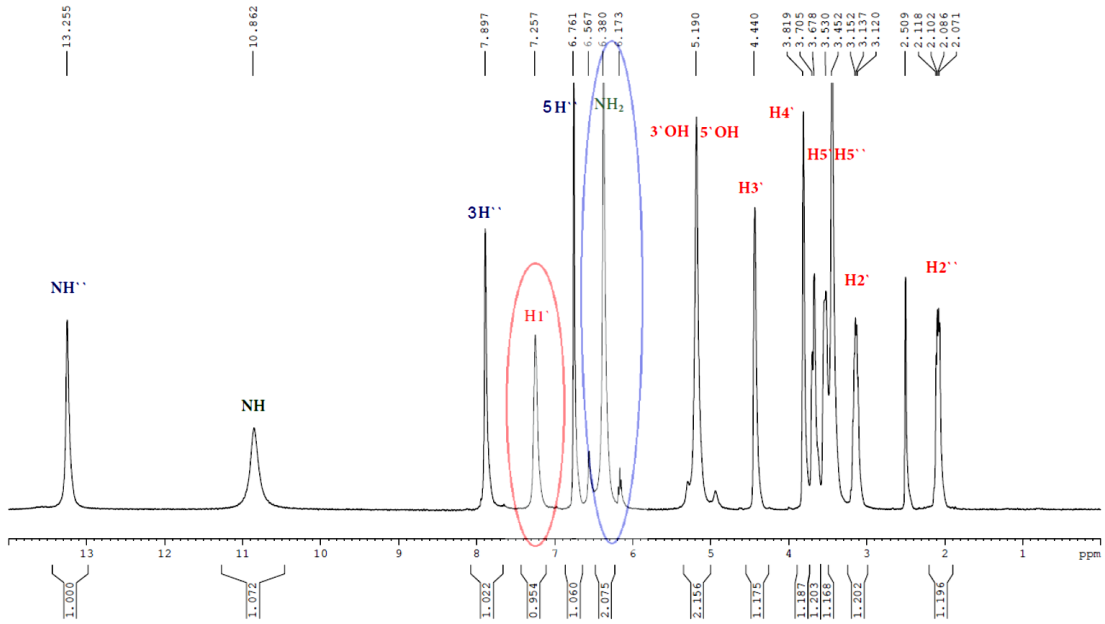


Figure 3.6: *[1]H NMR of 4-pyrazole-dG, displaying a 1ppm movement of the H1` peak.*

The identity of the H1` peak has been confirmed using correlation spectroscopy (cosy) NMR(Figure 3.7), where the H1`peak can be seen coupling with the H2` axial and equatorial protons, circled in **red**. Similarly, the heteronuclear multiple-bond correlation spectroscopy (HMBC) (Figure 3.8) displays the H1`peak coupling to the C1` carbon also circled in **red**.

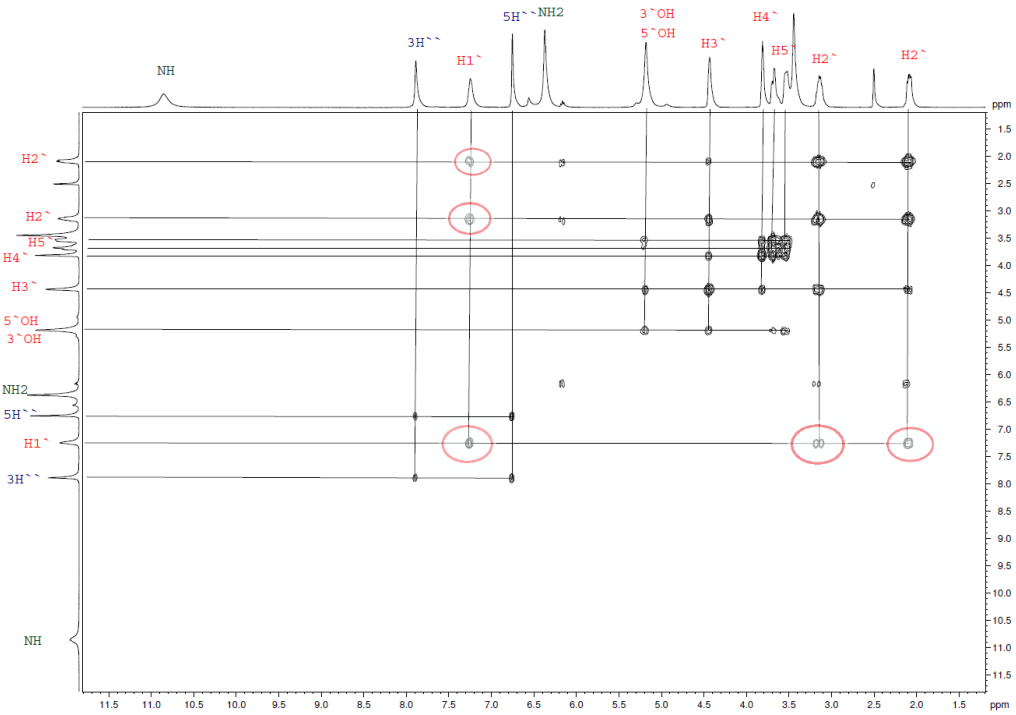


Figure 3.7 *Correlation spectroscopy of 4-pyrazole-dG, circled in red is the cross-coupling of the H1` peak with the axial and equatorial H2` peaks.*

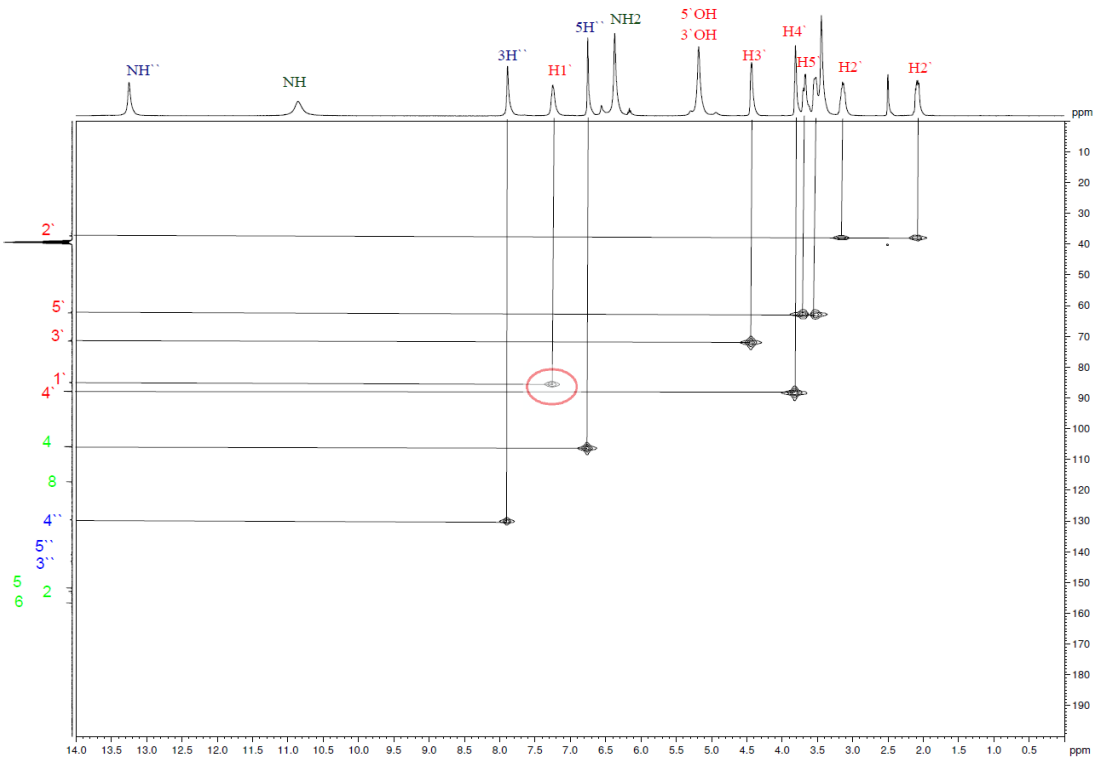


Figure 3.8: *Heteronuclear multiple-bond correlation spectroscopy of 4-pyrazole-dG, circled in red with the coupling of the H1` peak with the C1` peak.*

The cause for this phenomenon has been attributed to the lone pair of electrons on the nitrogen in the 5`` position of the pyrazole ring resulting in an increased electron density within the pyrazole ring. The resulting electronic resonance may have resulted in a deshielding interaction between the 3``-hydrogen and 1`-hydrogen. To further corroborate the hypothesis, a crystal structure of 4-pyrazole-dG was obtained and is displayed below in Figure 3.9.

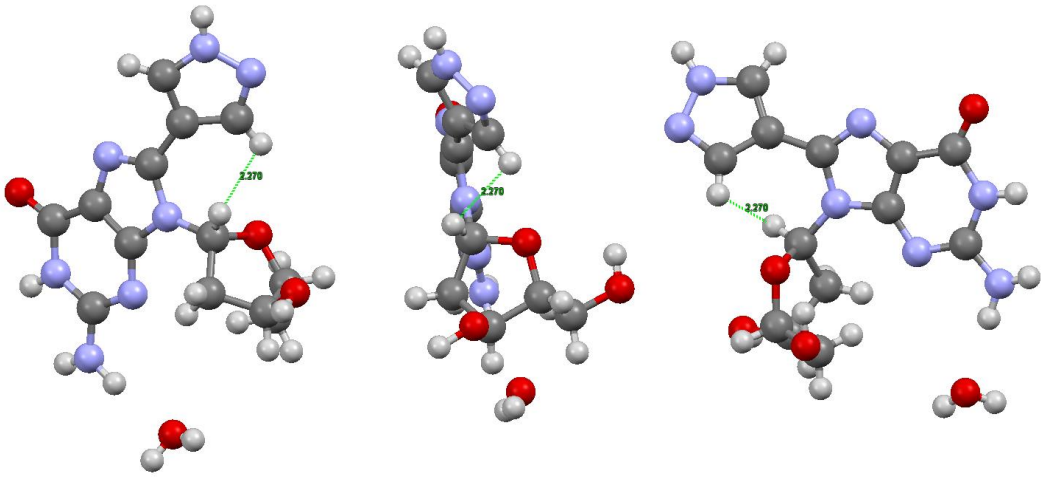


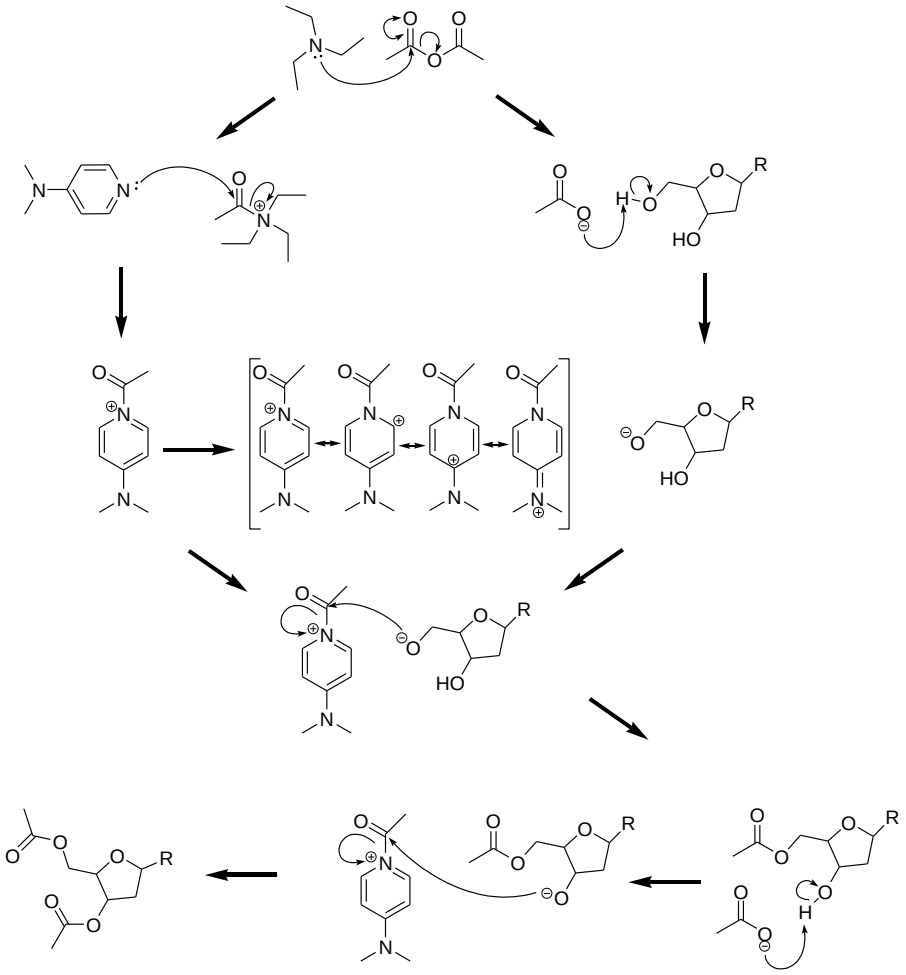
Figure 3.9 *Crystal X-ray structure of 4-pyrazole-dG*

The crystal X-ray structure shows that the distance between the 1`-H of the furanose ring and the 3``-H of the pyrazole ring are only 2.2 angstroms apart, a relatively short distance. From the

X-ray crystal structure it is evident that the furanose ring has adapted a *pseudo-trans* conformation. Although the deoxyribose can freely flip between *cis* and *trans* positions on 2'-deoxyguanosine, reports have shown that replacing the H^8 proton with bulky substituents (i.e. aromatic group) will orient the sugar ring into the *cis* formation only. This X-ray structure and the 2D NMR data provides plausible evidence that the 1`-H proton is placed within the local magnetic field of the pyrazole ring, resulting in the 1ppm de-shielding observed in the ^{1}H NMR as shown in Figure 3.6.

3.3.4 Modified Yamaguchi esterification

The Yamaguchi esterification reaction was discovered in 1979 by Masaru Yamaguchi *et al*.[65] The original reaction was used to functionalize aliphatic carboxylic acids with 2,4,6-trichlorobenzoyl chloride and 4-dimethylaminopyridine (DMAP), which works as a catalyst to form anhydrides. However, the scope of the reaction has improved in subsequent years. The reaction can now be used to functionalize alcohols using triethylamine (TEA) in place of 2,4,6-trichlorobenzoyl chloride. The reaction begins on the attack of the acetic anhydride by the triethylamine base. The carbonyl anionic species then deprotonates the primary alcohol, transferring the negative charge to the sugar. At the same time the DMAP attacks the carbonyl-triethylamine species, transferring the charge to the carbonyl-DMAP species. This reaction is driven towards the target ester product as the DMAP can stabilise the charge through resonance delocalisation as shown in scheme 3.23.



Scheme 23: *Mechanism of modified Yamaguchi esterification.*

Acetylation of the 5` group on the furanose sugar of derivatives of 2`-deoxyguanosine was been by reported Gubala *et al*[54]. The purpose of the acetylation of the library of compounds was to increase their solubility in chloroform, in order to demonstrate the assembly of G-quadruplex of the compounds by titrimetric addition of potassium iodide. The esterification reaction is highly sensitive to moisture and, accordingly, the reactions were carried out under nitrogen with anhydrous CH_3CN. It was reasoned that while suspended the reaction would proceed more slowly but would still progress as the target acetylated compounds would dissolve in solution thereby continually driving the reaction equilibrium toward the target product. However, thin layer chromatography showed the formation of no new products, as only starting material was observed for the attempted esterification reactions of compounds 3-hydroxy-dG, 4-hydroxy-dG, 4-pyrazole-dG and 5-pyrazole-dG, respectively. However, the most likely cause of the failure to synthesise the target products is the hydrolysis of the acetic anhydride to acetic acid due to the presence of moisture inside the reaction vessel.

3.3.4 Covalently linked G tetrad

Synthesis of the CLGT was dependent on the synthesis of the 5`-azide compound. The proposed additional Huisgen cycloaddition click chemistry reaction step was reported to be reliable and highly reproducible[66]. However, functionalization of the 5`-OH proved highly problematic. The synthesis of a variety of leaving group intermediate compounds was attempted, (Reactions 3, 4 and 5) However, none of the target compounds could be successfully synthesised. A potential cause for this may have been hydrogen bonding between the 5`alcohol and the -NH_2.

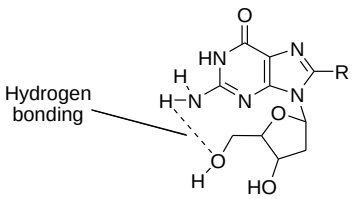
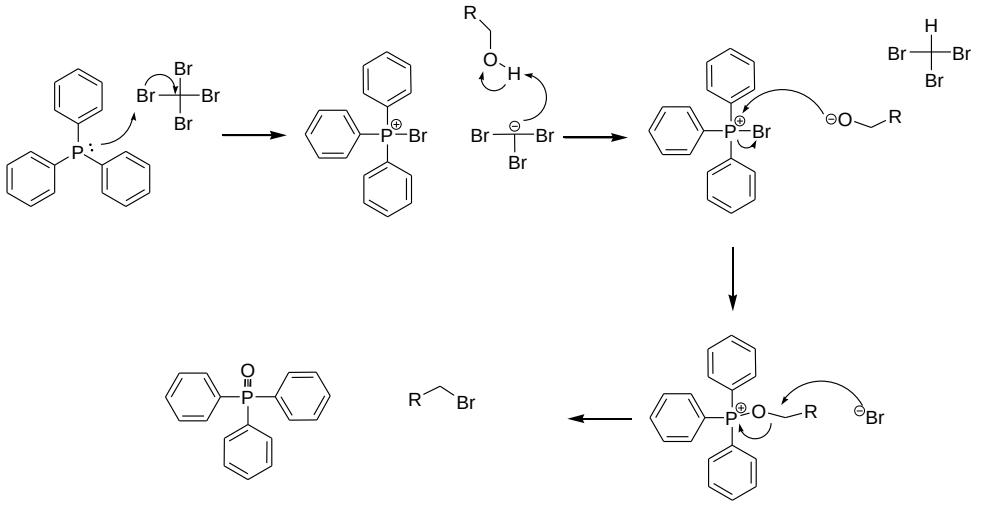


Figure 3.10: *Potential site of hydrogen bonding between the 5` alcohol and the NH_2 of the base.*

However, this leaves the question as to why the 3`OH was not substituted for the various reaction products. The 3`OH is a secondary alcohol, which leaves it less susceptible to S_N2 substitution. It may be that the 5` and 3` alcohols are partially sterically blocked by functionalization of an aryl compound in the C^8 position.

3.3.5 Appel reaction

The Appel reaction was devised by Rolf Appel in 1975[67]. The Appel reaction converts alcohols into alkylhalides using triphenylphosphine (TPP) with carbon tetrachloride, carbon tetrabromide or carbon tetraiodide. The mechanism of the Appel reaction begins with formation of the phosphonium salt by the nucleophilic addition of the halide to the TPP. The alcohol is deprotonated by the resulting carboanion to produce a methane trihalide species and an alkoxide anion. The alkoxide anion nucleophilicly displaces the halide followed immediately by an S_N2 reaction of the liberated halide nucleophile. The resulting products are the desired alkyl halide species and triphenylphosphine oxide. The Appel reaction selectively favours primary alcohols over secondary alcohols, which in turn are favoured over tertiary alcohols.



Scheme 3.24: *Mechanism of the Appel reaction*

The reaction was performed twice as the Appel reaction selectively favours primary alcohols over secondary alcohols. The first reaction was performed using 1.1 equivalents of the carbon tetrabromide and TPP reagents, relative to the quantity of (3-Ac-dG). It was assumed that the reaction would preferentially brominate the 5` primary alcohol in place of the secondary 3` alcohol. However, the first reaction yielded only starting material. Consequently, the reaction was repeated using 2.3 equivalents of carbon tetrabromide and TPP. The crude product showed no reaction as confirmed by ^{1}H NMR analysis. The ^{1}H NMR is attached in the appendix.

3.3.6 Conversion to mesyl or tosyl leaving group

Mesyl and tosyl groups are reported as good leaving groups for S_N2 substitution reactions[59] and generally performed in pyridine at room temperature[68]. Unfortunately at this temperature, the crude products were analysed by ^{1}H NMR and revealed no reaction had taken place. A literature search revealed a publication of a successful mesylation on a target alcohol at an elevated temperature[69]. Therefore, the reaction was repeated at 50 ˚C. The products of the repeat reaction were also analysed by ^{1}H NMR, and did not show the formation of target product. It was reasoned that the pK_a of the pyridine was too low to deprotonate the 5` alcohol. The reaction was repeated using a stronger base called triethylamine. Unfortunately, ^{1}H NMR analysis of the crude product revealed that the desired product was not formed. A possible alternative was provided in a recent publication showing that the successful mesylation of the target alcohol requires the reaction to be performed at -30 ˚C[70]. Due to time constraints, the reaction was not repeated at -30 ˚C.

Chapter 4

Testing of 8ArdG derivatives with telomeric DNA in novel assay and Dual Polarisation Interferometry

4.1 Chapter Introduction

4.1.1 Telomeric DNA depositions on zeonor and gold slides as a novel screening methodology for the formation of GQs

The second aim of the project was to use the library of compounds synthesised in Section 3, as test samples in the development of a novel screening process to detect small molecules capable of binding to or inducing the formation of G-quadruplexes. Since the formation of G-quadruplex structures is known to be enthalpy-driven, the equilibrium constant is temperature-dependant[71]. As a result, studies performed generally utilise techniques where the measurement is based on structural changes of GQs as a function of increasing temperature. Circular dichromism[72], NMR[73], UV melting profiles[74] and differential scanning calorimetry[75] are typical examples of methods used to probe the interference of molecules with G-quadruplex DNA structures. Significant advances have been also made to characterize the structure of surface-bound DNA probes using optical and contact methods such as atomic force microsocopy[76], x-ray photoelectron spectroscopy$_{77}$ and surface plasmon resonance[78]. However, none of the techniques mentioned so far are suitable for high-throughput screening of large numbers of samples. Also, their high initial price along with high running cost and demanding maintenance makes their widespread use in this field limited. The novel technique proposed for this research was an optical sensor based on the fluorescence quenching phenomenon of folded, labelled DNA on gold surface, as illustrated in Figure 4.1.

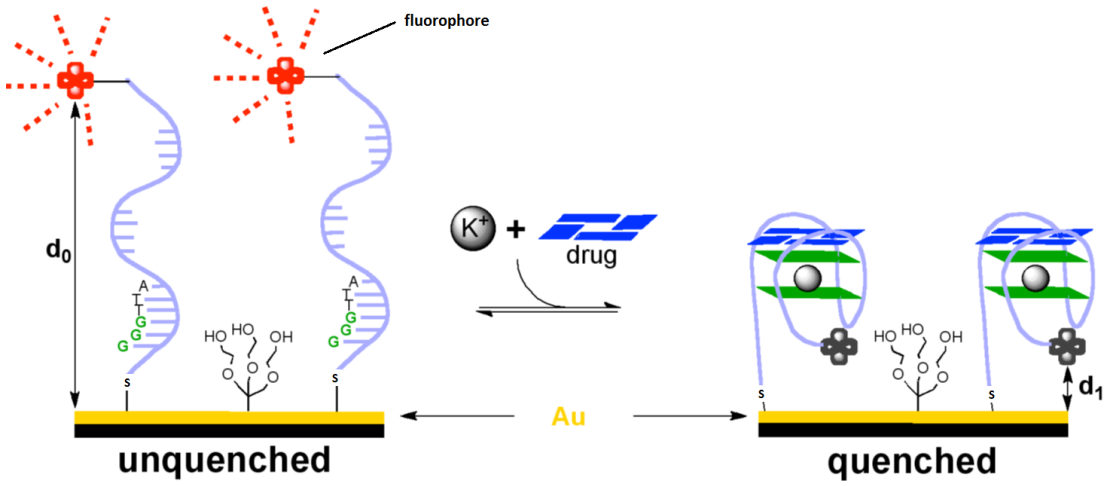


Figure 4.1: *On the left, fluorescently tagged telomeric DNA is covalently bound to a gold surface by a sulphur-gold bond. After the addition of a small molecule, the telomeric DNA undergoes G-Quadruplex folding, drawing the fluorophore to the gold surface and thus quenching the fluorescent signal.*

The principle of the fluorescence quenching method is shown in Figure 4.1. In this simple illustration, human telomere (repetitive motif = TTGGG), modified with a thiol group at the 5'

end and cyanine 5 (Cy5) at the 3' end was deposited on a gold surface, and covalently bound to the gold surface by the enthalpy driven formation of a covalent sulfur gold bond. The proposed novel screening methodology was to be performed by first measuring the signal of the florescent Cy5 attached to DNA bound to the gold surface without the addition of any small molecules, while the DNA was standing perpendicular to the gold surface. Initially the Cy5 fluorophore and the gold surface were to be as far away from each other as the length of the DNA could allow. At this distance the gold surface would have no impact on the fluorescent signal. This distance is represented as $\mathbf{d_0}$ (as shown in Figure 4.1). The unquenched fluorescent signal was to be recorded and measured on a Perkin Elmer fluorescent scanner. Then one of the novel compounds synthesised in section 3.1, as part of the library of compounds, was to be dissolved in H_2O with potassium iodide and deposited onto the DNA bound to the surface of the gold slide. Then the fluorescent signal was to be measured again. The strategy behind this novel process was that when the immobilised DNA came in contact with the novel compound the telomeric DNA would fold into a G-quadruplex. When the telomeric DNA is folded, it would decrease the distance between the Cy5 fluorophore the gold surface to $\mathbf{d_1}$, so that when the fluorophore is within 5nm of the gold surface, the fluorescent signal is quenched[79]. The novel methodology was intended to measure the decrease or disappearance of the fluorescent signal as the indicator for the formation of G-quadruplex. All DNA depositions were performed using a computer controlled automated deposition that can deposit 0.5 nL of sample solution per droplet. The depositions were performed using a varying numbers of droplets of 0.5 nL per individual spot using 1 μM concentration of the oligonucleotide solution.

4.2 Depositions of telomeric DNA on gold and Zeonor slides

The analysis of the ability of small molecules to induce GQ formation in telomeric DNA was performed using a Perkin Elmer fluorescent scanner. The scanner was loaded with sample slides of the dimensions 25 × 75 × 1.1mm. The DNA that was used in the deposition consisted

of telomeric 30-mer strands capped with a Cy5 fluorophore and a thiol group on the terminal ends. The gold slides used were borosilicate glass coated with a uniform 30-nm gold layer. In order to optimise the experiments, the initial trial depositions were performed on cyclo olefin polymer slides called Zeonor. The Zeonor slides were oxidised using oxygen plasma irradiation to enable the thiolated telomeric DNA to bind to the surface of the Zeonor slide *via* the formation of sulfur-oxygen covalent bonds.

Alongside the thiolated-Cy5-labelled telomeric 30-mer DNA, a control oligonucleotide was added for all deposition experiments. The control nucleic acid was 15-mer non-telomeric DNA molecule of the following sequence ACGGCAGTGTTTAGC, and with modification at the 5' end with Cy5 and 3' end with thiol group. The purpose of the control was to ensure that any resulting change in fluorescent signal, after the addition of the synthesised 8-Aryl-dG derivatives, could be attributed to the G-quadruplex formation and no other unexpected phenomena. In order to clearly distinguish between the telomeric DNA arrays and the control arrays, a distinctive deposition pattern was used to distinguish the thiolated, Cy5-labelled, telomeric 30-mer DNA from the control DNA. The sample DNA and control DNA were deposited on the surface of the Zeonor and gold slide as two different arrays that are clearly identifiable by the naked eye. The two arrays are displayed below (Figure 4.2).

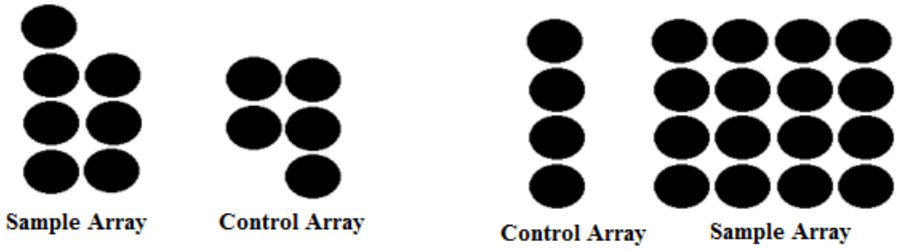


Figure 4.2: *Diagram illustrating arrays for deposition experiments 1 and 2 on the left and for deposition experiments 3 (A-F) and 4 on the right.*

4.2.1 Uniform dispersion of the deposition arrays

In order to ensure that the rate of evaporation did not affect the uniformity of the DNA dispersal on a surface, a deposition experiment was performed on a Zeonor slide, with the 30-mer telomeric DNA co-diluted with a high boiling-point diluent. The three compounds for co-dilution with water and the Cy5-labelled, thiolated, telomeric 30-mer DNA were: trehalose[80], glycerol and dimethyl sulfoxide (DMSO). The 3 compounds were selected because they would

not interact with the telomeric DNA or the novel 8-Aryl-dG derivatives and they would increase the surface tension of the water, thus reducing the rate of evaporation. The three reagents were added to the 10µM DNA as a 2% (w/v) solution making up samples with 5 and 1 µM DNA. The six solutions were deposited on an oxidised Zeonor slide at 20 droplets per array. The DNA was allowed to incubate for 2 hours and was washed with 1% (w/v) solution of sodium dodecyl sulfate (SDS). SDS is an anionic surfactant and its purpose in DNA washing is to remove any unbound DNA from the surface of the chip. This ensures that only the DNA that is covalently bound to the surface of the chip will remain on the surface, and be available to produce a fluorescent signal. Washing also prevents any unbound DNA competing with bound DNA for interaction with small molecules to form a G-quadruplex. The results of the optimisation experiment demonstrated that the most uniform depositions were 1 µM Cy5-labelled, thiolated, telomeric 30-mer DNA with 2% (w/v) (52.86 mM) trehalose at 20 drops per array. The results are displayed in figure 4.3 below.

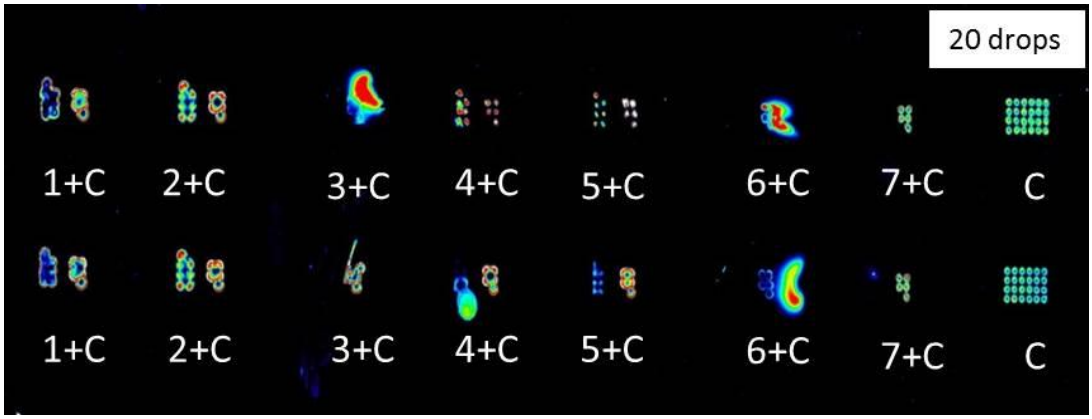


Figure 4.3: *30-mer, thiolated, Cy5-labelled, telomeric DNA and control DNA (C), 1 µM 15-mer, thiolated, Cy5-labelled, non-telomeric DNA of sequence ACGGCAGTGTTTAGC were deposited on an oxidised Zeonor slide at varying concentrations using 20 drops per array. 7 solutions of DNA were deposited on the Zeonor surface. The first solution deposited on the surface of the slide was untreated 10 µM 30-mer DNA in H_2O (1), two solutions in H_2O, 5 µM 30-mer DNA and 5 µM trehalose (2) and 1 µM 30-mer DNA with 1 µM trehalose (5) were deposited on the surface of slide. Glycose in aqueous solutions of 5 µM and 5 µM DNA (3) and 1 µM and 1 µM (6) were also deposited on the surface of the slide. The last 2 solutions deposited on the slide surface were 5 DMSO 5 µM with 5 µM DNA (4) and DMSO 1 µM and 1 µM DNA (7).*

4.2.1 Deposition of the oligonucleotides on gold slides

After it was confirmed that fluorescently tagged DNA could be successfully deposited onto the surface of a Zeonor slide and its fluorescent signal could be measured, a deposition experiment was performed on a borosilicate glass chip with a uniform 30 nm gold layer adhered to the surface. The gold slide was washed in piranha solution to remove any particles or impurities

that may have adhered to the surface. Piranha solution is a 3:1 mixture of concentrated sulphuric acid 98.0% (w/v) and hydrogen peroxide 30% (w/v) that is a strong oxidising agent that can remove any impurities that may be covalently bound to the surface of the gold. 1 µM 30-mer telomeric DNA in a solution of 2% (w/v) of trehalose (52.86 mM) in H_2O was deposited on the surface of the gold slide in arrays of 20 droplets per spot. After washing with SDS for 15 min, no fluorescent signal was observed on the gold surface. However, on the right-hand side of the image it is possible to observe a fluorescent emission in the same orientation as the deposition arrangement array. The purchased gold-coated glass substrates were of standard microscope slide dimensions (75mm × 25mm). The 30 nm layer of gold was coated on an area of 55 mm × 25 mm, leaving one part of the slide uncoated. The oligonucleotides were also deposited on this part of the slide and the observed single line of fluorescent signal corresponds to this part only (Figure 4.4).

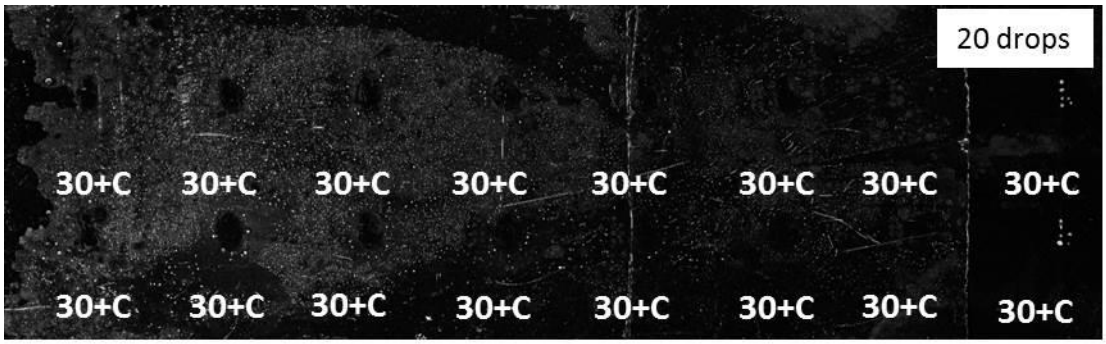
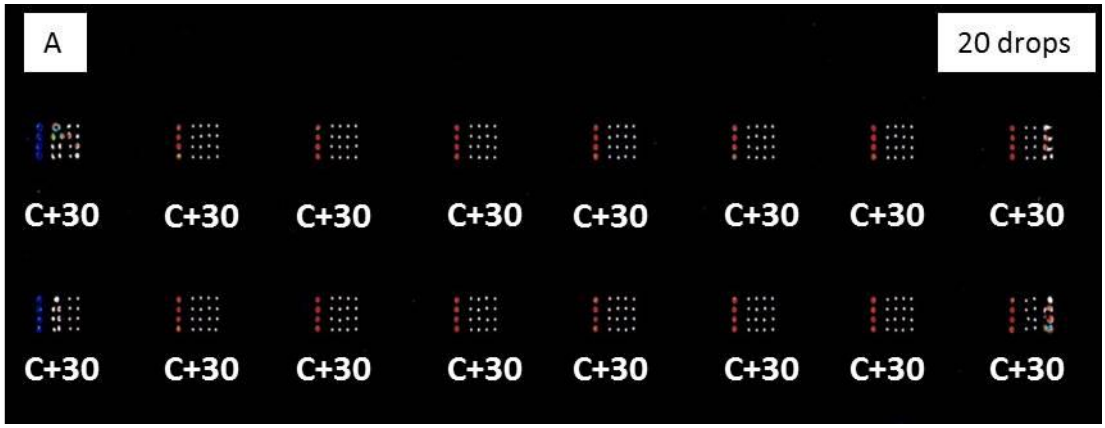


Figure 4.4: *30 nm gold slide was washed with piranha solution. 1 µM 30-mer, thiolated, Cy5-labelled, telomeric DNA with 1 µM trehalose displayed as (30) and control DNA, 1 µM 15-mer, thiolated, Cy5-labelled, non-telomeric DNA of sequence ACGGCAGTGTTTAGC displayed as (C) were deposited on the slide after washing with piranha solution, After DNA deposition the slide was washed in a 0.1% (w/v) SDS solution.*

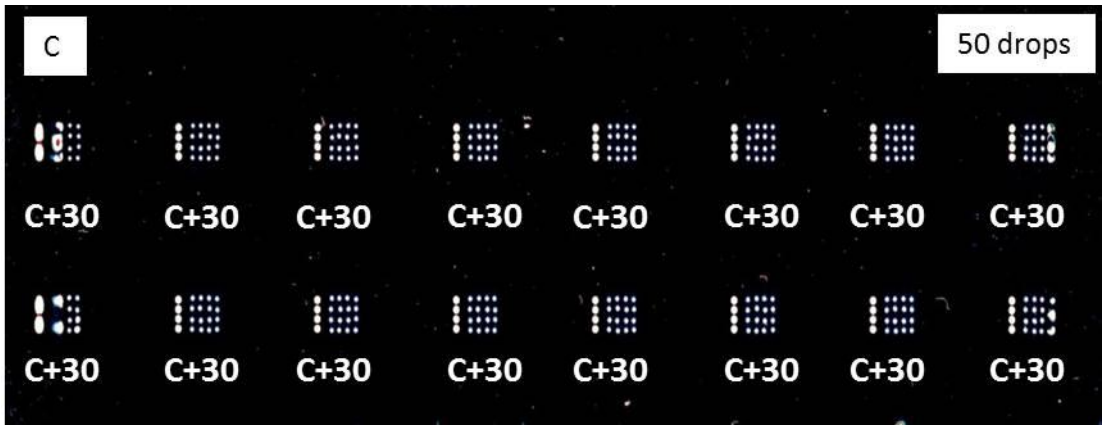
Two rationales were proposed to explain this observation. First, the DNA did not adsorb on the gold surface efficiently and was completely washed in the washing step with 0.1% (w/v) SDS solution. Second, the deposited, dried DNA molecules adapted 'flat-on' orientation on the substrate surface, thus rendering the Cy5 label close to the gold layer. In this case, the emission of Cy5 molecules would be effectively quenched by the underlying gold and no fluorescence signal would be observed.

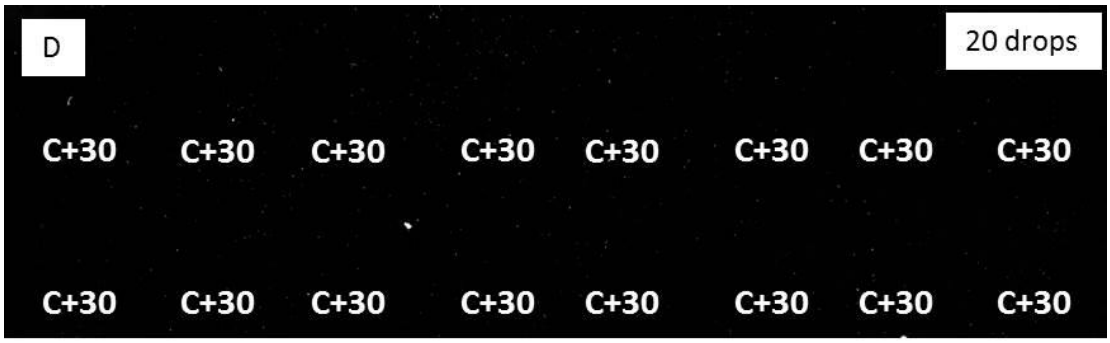
The experiment was repeated on brand new gold slides. New gold slides were also used to ensure that the piranha solution had not oxidised the gold film on the slide and disrupted the uniformity of the gold monolayer. However, the experiment was also repeated with different numbers of deposition droplets to observe if an increase in the quantity of DNA on the surface of the gold slide would have an impact on the fluorescent signal. The deposition of 1 µM Cy5-

labelled, thiolated, telomeric 30-mer DNA with 2% (w/v) (52.86mM) trehalose was performed on three separate gold slides. The three slides had a different number of drops per spot. The number of drops in each spot was 20, 30 and 50 respectively. The DNA was allowed to incubate for 1 hour. The fluorescent signal was measured before and after the slides were washed with 1% (w/v) solution of SDS, and the resulting images are displayed in figure 4.5 below.


A
20 drops
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30


B
30 drops
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30


C
50 drops
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30


D
20 drops
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30
C+30 C+30 C+30 C+30 C+30 C+30 C+30 C+30

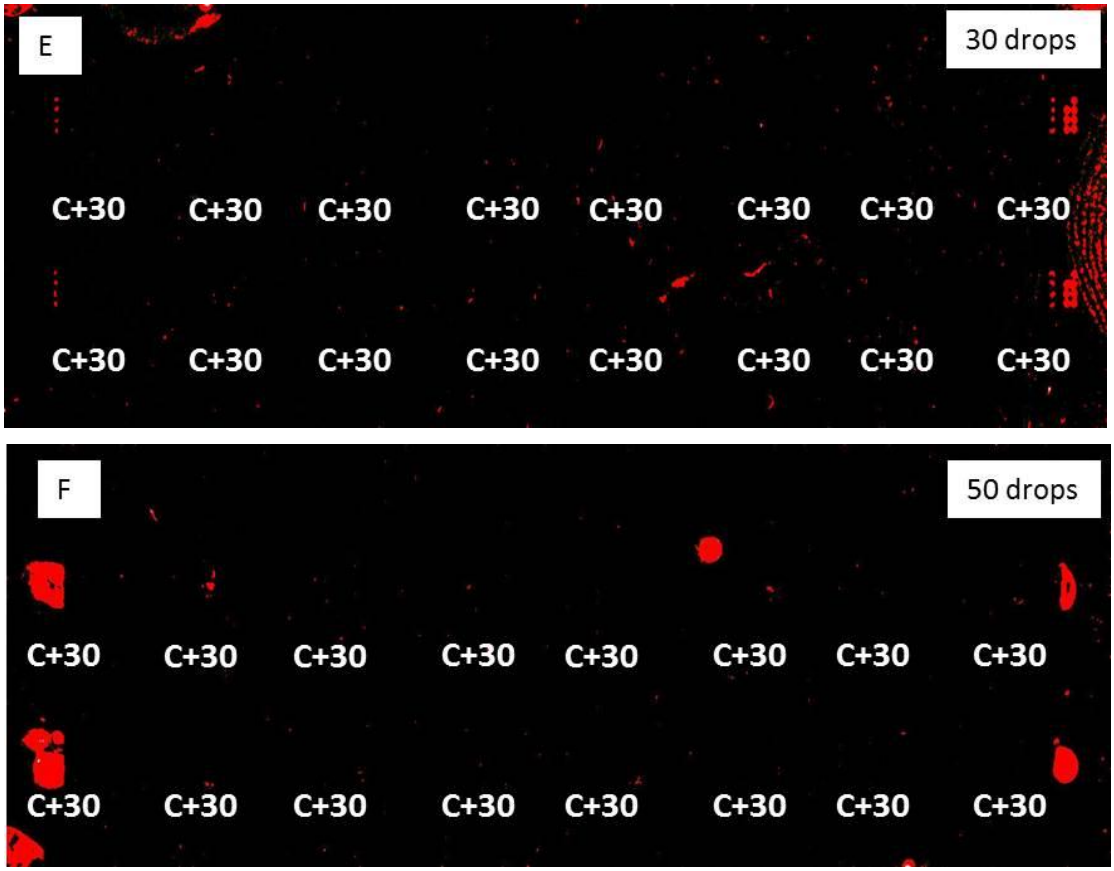


Figure 4.5: *The oligonucleotide deposition experiments. (A-F), depositing 1 µM 30-mer, thiolated, Cy5-labelled, telomeric DNA displayed as (30) and control DNA, 1 µM 15-mer, thiolated, Cy5-labelled, non-telomeric DNA of sequence ACGGCAGTGTTTAGC displayed as (C) on 30 nm gold slides. Experiments A, B and C are the fluorescent images of the slides before they were washed in an aqueous solution of SDS 2% (w/v), the number of droplets deposited per slide was 20 for A, 30 for slide B and slide C was deposited with 50 drops. Experiments D, E and F are the fluorescent images of the slides A, B and C after they were washed in an aqueous solution of SDS 2% (w/v).*

However, when the slides were washed, no fluorescent signal was detected, apart from slide E, where a clear fluorescence signal was observed on the glass (not gold coated) part of the slide. The fluorescence images were acquired on dried slides. From our observations two potential outcomes were deduced. The first deduction was that the DNA had failed to form a covalent bond between the thiolated terminus and the gold surface, and had been washed from the surface of the slide. For our second deduction, we hypothesised that if the DNA bound to the surface of the gold had adapted flat-on orientation, in which the signal was effectively quenched by the gold layer, the addition of the 8-Aryl-dGs with potassium iodide could induce the formation of GQs. This in turn could rearrange the orientation of the oligonucleotide on the surface, thus pushing the Cy5 dye away from the quenching gold. In order to test our second deduction, we have therefore created individual wells (using a silicone gasket) on the deposited

slide and in each well we have incubated 100 uM of the synthesised 8-Aryl-dG derivatives along with 100 uM of KI. The gasket provided individual wells on the surface of the gold slide with more than 100 uL volume capacity. The concept is illustrated on figure 4.6 below.

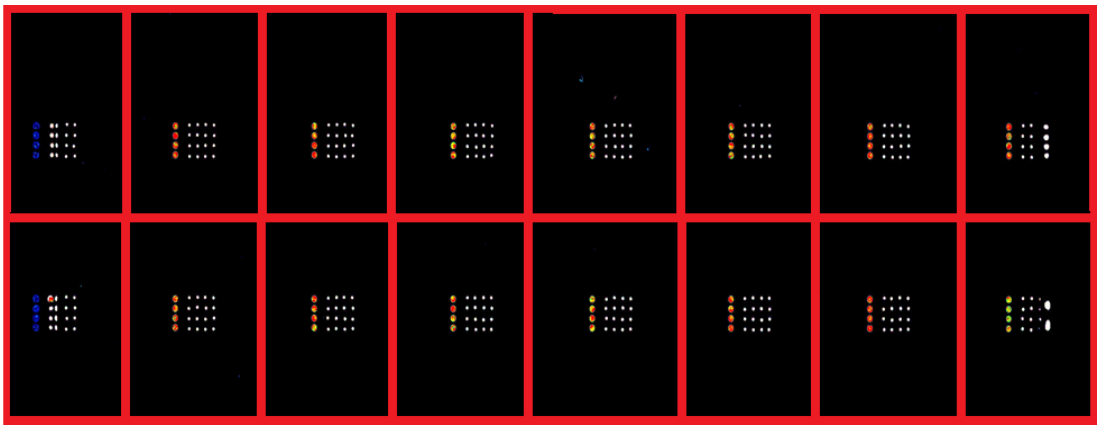


Figure 4.6: *Diagram of a silicone gasket placed on the surface of a 30 nm gold slide with 1 µM 30-mer, thiolated, Cy5-labelled, telomeric DNA deposited on the surface. The red lines are the gasket walls, the areas inside the red boxes are the wells that can contain 100 µL of solutions of the Aryl-2`-deoxyguanosine derivatives, to incubate on the deposited telomeric DNA.*

The gold slides were allowed to incubate with the gasket in the solution of prepared 8ArdG analogues for 2 hours. The solutions were then carefully removed with the use of a vacuum line. The slides were allowed to air dry, while protected from any source of light to prevent photo bleaching of the dye. The fluorescence of the slides was re-measured again and the resulting image is provided on figure 4.7.

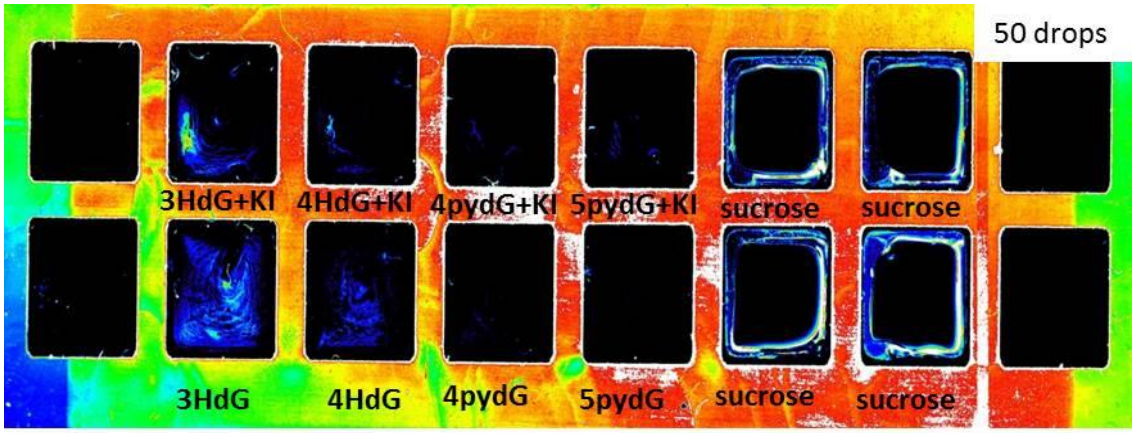


Figure 4.7: *1 µM 30-mer, thiolated, Cy5-labelled, telomeric DNA and control DNA, 1 µM 15-mer, thiolated, Cy5-labelled, non-telomeric DNA of sequence ACGGCAGTGTTTAGC were deposited on the 30 nm gold slide using 50 drops per array. A gasket was adhered to the surface of the slide and four solutions of acetonitrile and water in a 1:1 ratio, containing 100 µM of compounds 4-pyrazole-dG (4pydG), 5-pyrazole-dG (5pydG), 3-hydroxy-dG (3HdG) and 4-hydroxy-dG (4HdG) were added to the wells and allowed to incubate for 1 hour. Four additional solutions (4pydG+KI), (5pydG+KI), (3HdG+KI) and (4HdG+KI) containing the same*

respective Aryl compounds at 100 µM concentration with 100 µM of potassium iodide were also deposited into the wells on the surface of the gold slide and allowed to incubate for 1 hour. A solution of 100 µM sucrose in H_2O was also deposited in 4 of the wells as a control. The multi coloured pattern on the surface of the slide is residue from the silicone gaskets adhesive layer,that allows the gasket to stick to the gold slide and form water tight wells on the surface of the slide.

The fluorescence scan of the slide did not show any improvement, effectively, no signal was detected. The conclusions from these experiments support the first deduced hypothesis, that in fact, very little or no DNA was immobilized on the studied gold surface.

4.3 Dual polarisation interferometry

Dual Polarisation Interferometry (DPI) is an analytical technique that allows the real-time observation of molecular changes on a surface, thus DPI has been referred to as a molecular microscope. DPI uses polarised light from a laser passing down stacked waveguides. These waveguides are incorporated into the structure of a specialised Anachip composed of silicon oxynitride. Three channels are incorporated in the chip. Two of the channels are the sample channels through which the mobile phase and the sample compounds are injected. The third (middle channel) serves as a reference channel through which only the mobile phase passes. Sample molecules can adhere to the surface of an Anachip by both chemical and physical processes. The evanescent field emanating from the top waveguide can be analysed to display molecules bound to the surface of the Anachip in real-time. Changes in the resulting optical interference pattern are caused by changes in the structure and/or mass of any materials immobilised to the surface of the Anachip. These changes in the optical interference pattern can be analysed to deduce structural information, such as mass, thickness and density, about molecules adhering to the surface of the Anachip. DPI splits light from a laser into two waveguides. The first waveguide functions as a reference beam and travels through the centre of the Anachip. The second waveguide functions as the sensing waveguide and travels along the surface of the Anachip as illustrated in Figure 4.8. By combining the polarized light from the two separate waveguides, a two dimensional interference pattern is formed. DPI analysis rotates the polarized light emitted from the laser to alternatively excite two alternate polarization frequencies of the waveguides. The resulting interference fringes for both polarizations are analysed using a high performance charge coupled device (CCD) camera. The camera allows the instrument to calculate the mass of a material on the surface of the Anachip, accurate to 0.1 angstroms, and the density of the material, accurate to 0.1 picogram/mm^2.

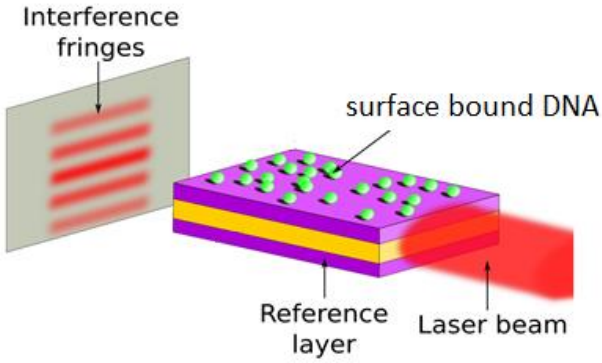


Figure 4.8: *Diagram illustrating polarised light passing through the Anachip, forming interference fringes on the far side, that are analysed by a camera to calculate changes in mass, thickness and density of a material bound to the surface of the Anachip.*

The polarization of light emitted from the laser is computer controlled, allowing rapid switching (10 Hz) of the polarization frequency. This rapid switching allows real-time measurements of chemical reactions taking place in the material on the surface of the Anachip.
DPI has been used to characterise biochemical interactions by quantifying any conformational change while measuring reaction rates, affinities and thermodynamic properties[81].
DPI analysis requires silicon oxynitride Anachips. However, depending on the chemical composition of the sample material, a wide variety of specialised Anachips with coated surfaces are available. The surface of differing Anachips are coated with different molecules that facilitate the adhesion of various materials. In this project, thiol and amine-coated Anachips were used.
DPI in many respects is similar to HPLC. The system utilizes a mobile phase which transports the molecules of study to the surface of the Anachip. The flow rate of the mobile phase is controlled by a pump which can be manually or computer controlled. The mobile phase is allowed to flow over the surface of the chip. DPI uses refractive index (RI) to calculate and measure changes to the surface of a chip. The analysis requires the Anachip to be calibrated against a standard known to the DPI software. The chip is calibrated against the mobile phase of PBS and two high purity standards, water and ethanol. The RI of water and ethanol are measured by the DPI software, and if the RI of the injected water and ethanol correspond to the pre-defined value known to the DPI software, then the calibration is successful.

4.3.2 DPI results

The library of compounds synthesised and described (section 3.1) consisting of the compounds 4-hydroxy-dG, 3-hydroxy-dG, 4-pyrazole-dG and 5-pyrazole-dG were studied for binding to GQ using DPI. The compounds were analysed using two methodologies. The first methodology was performed using Anachips that were coated with sulfhydryl groups (SH).The second methodology was performed using amine-coated Anachips (NH_2). Both methodologies were outlined in detail in section 2.4. The data obtained for the changes in thickness, mass and density to the bound telomeric DNA after the addition of 100 µM of the four compounds + 100 µM KI are displayed in section 4.3.2.

4.3.2.1 Thiol coated Anachip surface

The first methodology for DPI analysis was performed using thiolated Anachips. As the telomeric DNA possessed a thiol terminal group there was no need for a cross-linker to bind the DNA to the surface of the Anachip. The DNA binds to the surface of the Anachip via the formation of sulphur-sulphur covalent bond. In order to allow the bond to form, the flow rate of the instrument was stopped and the DNA was allowed to incubate on the surface of the

Anachip at room temperature. For DPI experiment (#1) the incubation time was 2 hours. Following the failure of the DNA to bind to the thiol surface in the DPI experiment #1 (table 4.1), a literature search was performed to determine why the thiolated telomeric DNA was not successfully binding to the thiol surface. A publication by Spain *et al*. revealed that 10 hours is the optimum time for the assembly of a DNA monolayer on a gold surface[82]. Accordingly the incubation time for thiol-coated experiments was changed to 10 hours.

DPI experiment #1*: 18 and 22-mer injection with 2 hour incubation time*

Thiolated 18-mer and 22-mer telomeric DNAs at a concentration of 100 µM were used. Because no binding was observed, as shown in table 4.1, it was decided to incubate a robust antibody (goat anti-human IgG) to observe any form of binding. As seen in table 4.1 the antibody was able to adsorb onto the Anachip surface in sufficient mass and density, and the measured thickness is in agreement to previously obtained data by the surface science group at BDI[83].

Table 4.1 DPI results for injection of 18 and 22-mer, telomeric DNA on thiol-coated Anachip with Control thiolated-protein

Name	**Thickness (nm)**	**Mass (ng/mm^2)**	**Density (g/cm^3)**
baseline	0	0	0
100uM 18-mer	0	0	0
100uM 22-mer	0	0	0
Control IgG protein	3.1	1.46	0.47

DPI experiment #2*: Injection of 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG compounds onto thiol surface*

A separate thiol-coated Anachip was calibrated and this time, 1µM 30-mer DNA modified with SH group on its 5' end was injected on the surface and allowed to incubate for 10 hours. Following the successful binding of the DNA to the thiol surface, solutions of acetonitrile and water in a 1:1 mixture, containing 100 µM of compounds 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG and 100 µM of potassium iodide were injected. After the injections of 8ArdGs, the thickness value of the thiol-coated Anachip dropped below the baseline value. A control solution of 5 mM sucrose in H_2O was injected and was absorbed onto the surface of the chip.

Table 4.2 Injection of 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG on thiolated Anachip surface after 10 hour incubation of telomeric DNA

Name	**Thickness (nm)**	**Mass (ng/mm^2)**	**Density (g/cm^3)**
Baseline	0	0	0
1µM 30-mer DNA	0.16	0.08	0.52
5mM KI	0.16	0.08	0.48
100µM 3-hydroxy-dG	below baseline	below baseline	below baseline
100µM 4-pyrazole-dG	below baseline	below baseline	below baseline
100µM 5-pyrazole-dG	below baseline	below baseline	below baseline
100µM 4-hydroxy-dG	below baseline	below baseline	below baseline
5mM sucrose	2.03	0.29	0.14

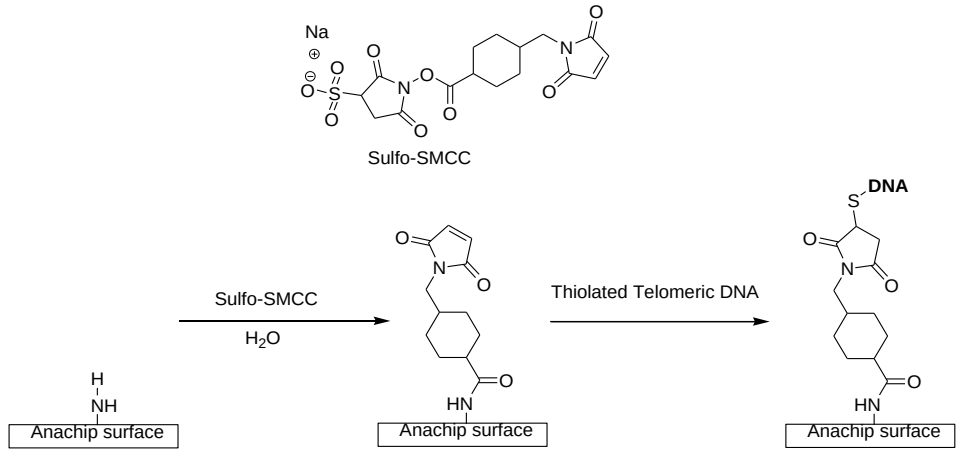
The DPI software used was Analight analysis version 1.7. The software is used to calculate changes in mass, thickness and density on the surface of a calibrated Anachip. The limitation of the software is that if the surface of the Anachip is not stable throughout the complete measurement, and some material is washed off the surface, the signal may drop below the original baseline level. In this case, the software will not be able to calculate the desired parameters and the outputs will be reported as errors. The important message that was learnt from this observation was that the chips will have to be calibrated and equilibrated in the solution for long time before the measurements.

4.3.2.2 Amine -coated Anachip surface

With the information about the software limitations in mind, we attempted to repeat the measurements again, leaving the chip incubating for a long period of time. However, because there was no thiolated Anachip available, an amino coated Anachip was used as an alternative. On the NH_2-coated Anachip, thiolated, Cy5-labelled telomeric 30-mer DNA was deposited As the DNA possessed a thiol terminal group, a linker compound was required to bind the DNA to the amine-functionalized surface. The linker used was sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC). Sulfo-SMCC is a water soluble cross linker used to cross-link thiol and amine groups[84]. It is comprised of an amine-reactive N-hydroxysuccinimide and a sulfhydryl-reactive maleimide group.

The N-hydroxysuccinimide ester reacts with primary amines at neutral to slightly alkaline pH (7-9) to form stable amide bonds. The maleimide functional group can be reduced by sulfhydryl groups at neutral pH (6.5-7.5) to form stable thioether bonds, as illustrated in scheme 4.3.1. Sulpho-SMCC must be prepared immediately prior to use as a linker, as the

malemide functional group can be hydrolysed to succinimide in the presence of water over time.



Scheme 4.1: *Sulfo-SMCC binding to an amine surface and serving as a DNA linker*

DPI experiment #3*: 4-pyrazole-dG deposited on an amine-coated Anachip*

DPI experiment #3 was performed on an amine coated Anachip. After the chip calibration, 5 mM of Sulfo-SMCC linker was dissolved in H_2O, and injected onto the amine coated surface. The sulfo-SMCC appeared to bind successfully to the surface of the Anachip, as evidenced by an increase in thickness from 0 to 1.0 nm. 10 µM of 30-mer thiolated telomeric DNA was injected onto the activated surface at a flow rate of 5 mL/min. The DNA was successfully bound to the linker compound. After injection of the DNA, the thickness on the surface of the chip increased to 1.54 nm, indicating a horizontal (flat on) orientation of the oligonucleotide. This phenomenon was observed before, especially on positively charged surfaces by senior researchers, Dr. Nam Cao Hoai Le and Dr. Vladimir Gubala at BDI (unpublished data). After the successful DNA immobilisation, 100 µM of 4-pyrazole-dG with 100 µM of potassium iodide in a 1:1 mixture of acetonitrile and water was injected onto the Anachip surface at 5 mL/min. The thickness of the chip surface decreased from 1.5 nm to 0.6 nm, indicating a subtle change in the DNA conformation. The results are summarized in table 4.3.

Table 4.3 DPI results for injection of 100 µM 4-pyrazole-dG on an amine-coated Anachip

Name	**Thickness (nm)**	**Mass (ng/mm^2)**	**Density (g/cm^3)**
Baseline	0	0	0
Sulfo-SMCC	1.0	0.48	0.46
30-mer telomeric DNA	1.5	0.83	0.53
4-pyrazole-dG	0.6	0.19	0.32
4-pyrazole-dG (2)	0.6	0.22	0.37

***DPI experiment #4**: 5-pyrazole-dG deposited on an amine-coated Anachip*

DPI experiment #4 was performed using the same protocol as the experiment #3.This time the thickness on the DNA layer on the Anachip surface was calculated as 1.1 nm. 100 µM of 5-pyrazole-dG with 100 µM of potassium iodide in a 1:1 mixture of acetonitrile and water was injected onto the Anachip surface at 5 mL/min. The Analight software showed the thickness of the chip surface decreased to 0.3 nm. The results are summarized in table 4.4.

Table 4.4 DPI results for injection of 100 µM 5-pyrazole-dG on amine-coated Anachip

Name	Thickness (nm)	Mass (ng/mm^2)	Density (g/cm^3)
Baseline	0	0	0
Sulfo-SMCC	1.0	0.61	0.64
30-mer telomeric DNA	1.1	0.52	0.50
5-pyrazole-DG	0.3	0.13	0.43

4.4 Discussion

4.4.1 DNA depositions of Cy5-thiolated telomeric 30-mer DNA on gold slides.

As previously stated (section 4.4.1.3), the optimum incubation time for a DNA monolayer to form on a surface was found to be 10 hours. However, the initial DNA depositions were performed with droplets of 0.5nL in which the solution evaporated quickly in approximately 20 min. Despite the relatively short incubation time in solution, the DNA depositions on gold slides using telomeric DNA usually proceed with successful results. However, washing the gold slides in aqueous 2% (w/v) SDS seemed to completely remove the fluorescent material from the surface of the gold chip. In solution, it is possible that the DNA molecules adopt a vertical (upright) position as illustrated (figure 4.9 below). In this orientation the Cy5 fluorophore is far away from the gold surface, which allows the unhindered fluorescent emission.

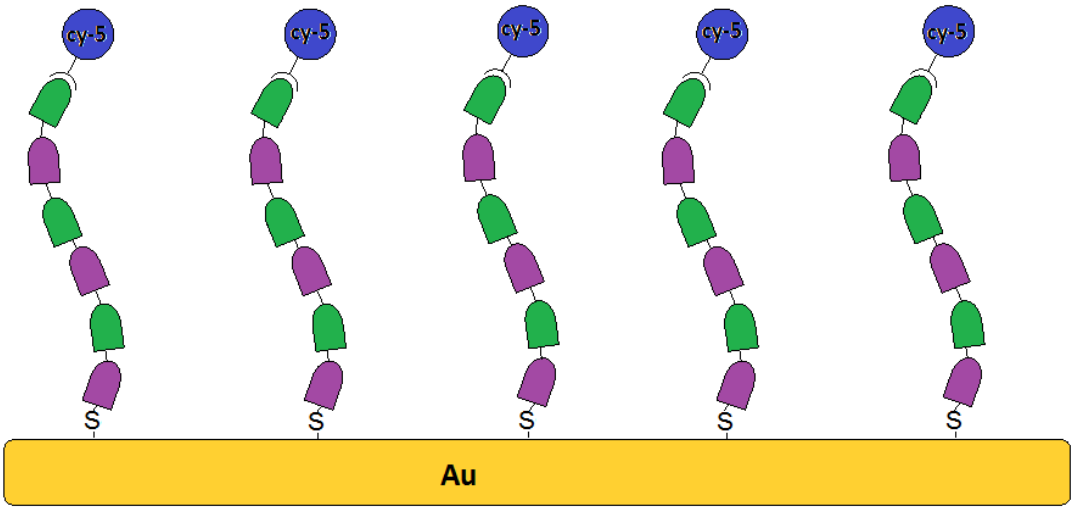


Figure 4.9: *Cy5-labelled 30-mer telomeric DNA bound to a gold surface in an "upright" orientation.*

However, when the DNA was washed with a surfactant solution and thoroughly dried in the stream of nitrogen, it is possible that the DNA layer was compacted into a flat-on position, as illustrated on figure 4.10. This may have brought the Cy5 into close proximity to the gold surface, thus effectively quenching the fluorescent emission.

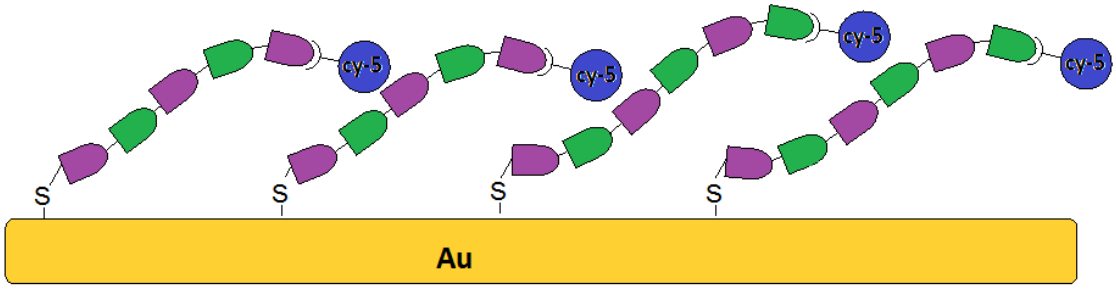


Figure 4.10: *30-mer, Cy5-labelled, telomeric DNA, bound to a gold surface, washed flat. The Cy5 fluorophores are in close proximity to the gold, quenching the fluorescent emission.*

Another potential cause for the disappearance of the fluorescent signal on the gold slides after washing may be that a large amount of DNA was non-specifically bound to the surface, forming multilayers. The detected fluorescent emission detected before the slides were washed may have originated from non-covalently bound DNA that was deposited on top of the covalently bound DNA lying across the gold surface, as illustrated in Figure 4.11.

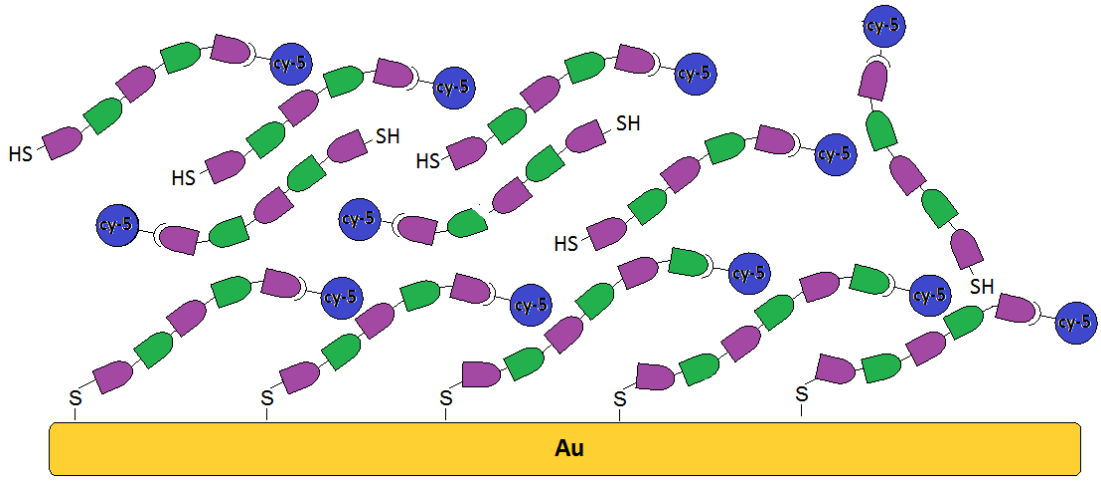


Figure 4.11: *Non-covalently bound, disordered DNA, deposited on the surface of the covalently bound DNA.*

4.4.1.1 Addition of compounds to disrupt gold quenching of Cy5 fluorophore

It was reasoned that the addition of the 8ArdG could induce the formation of a GQ and a change in fluorescence signal could have been observed. The idealised mechanism is illustrated on figure 4.12. Unfortunately, no signal was measured. The results of this experiment lead to conclusions that it is very likely that no DNA was immobilised on the gold surface. Because this type of reaction was previously reported in many scientific articles, the inability to attach thiolated DNA molecules to gold coated surface was attributed to the methodology of deposition using the sciFLEXARRAYER S3 instrument. An alternative methodology, more likely to successfully bind the DNA to the gold surface is outlined in the proposed future work, detailed in section 5.3.

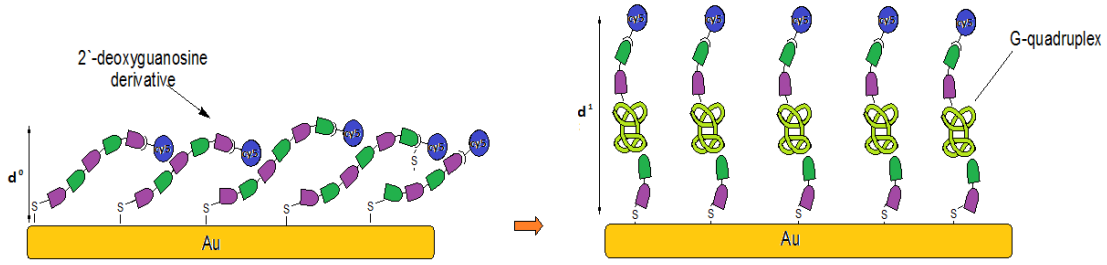


Figure 4.12: *Addition of Aryl-2`-deoxyguanosine derivative to non-covalently bound telomeric DNA, The distance between the Cy5 fluorophore and the gold surface before the addition of Aryl-2`-deoxyguanosine derivative is represented as (d^0). The formation of GQ, in the telomeric DNA, increases the distance (d^1) between the Cy5 fluorophore and the gold surface, eliminating the quenching effect.*

4.4.2 Dual polarisation Interferometry

4.4.1.2 DPI experiment 1

The first DPI experiment was performed on a thiol-coated Anachip. The corresponding display data can be found in appendix 3, DPI experiment 1. The experiment was performed using 18-mer and 22-mer nucleic acids. Both were labelled with Cy5, and modified with a SH group on the 3'-end.

An additional experiment was performed to test if increased concentration of the DNA can result in successful immobilisation. This time, 100 µM of 18-mer and 22-mer thiolated DNA was incubated on the SH modified Anachip. Unfortunately, the results for the DNA injections were the same as in DPI experiment 1. To validate that the DPI hardware was functioning correctly a random concentration (10 µM) of goat anti-human IgG was injected. Goat anti-human IgG possesses useful functional groups such as NH_2, SH and COOH, some of which can form strong dative bonds with thiols on the Anachip surface. The rationale was to observe if the bulky immunoglobulin molecule could be adsorbed on the surface, thus confirming that the hardware of the instrument was working properly. The DPI display analysis showed a 2.6 nm increase in thickness upon injection of the antibody. The calculated density, mass and thickness change was in agreement with previously reported values by senior BDI researchers Dr. Nam Cao Hoai Le and Dr. Vladimir Gubala (unpublished data). This indicated that the analysis system was working properly and the inability to immobilise the DNA on the thiolated surface was related to the DNA itself or inadequate incubation time. In addition, thiolated DNA was received as a lyophilised sample, which needed to be reconstituted in appropriate buffer. It is a known phenomenon that two thiolated DNA molecules can form a dimer in solution, accelerated by the formation of –S-S- bonds. The DNA samples are in these cases often pre-treated with solutions of DTT (dithiothreitol), which effectively breaks disulphide bonds. However, no pre-treatment was performed with these samples before they were incubated on the thiolated Anachip surface.

4.4.1.3 DPI experiment 2

The second DPI experiment was performed with nearly identical conditions as the first one. This time, problems with data analysis were observed. The DPI display data did not show any increase in thickness on the thiol surface. The Analight 1.7 software failed to calculate the change in thickness, because the measured thickness was less than the baseline thickness value. As mentioned previously, the DPI software used to analyse the crude data has its limitations. The most significant is that the software is unable to calculate the results if some material is washed off the surface, and when the signal drops below the original baseline level. In this

case, the software outputs will be reported as an analyses calculation error. Despite that the crude data might have indicated a change in the observable signal (figure 4.13), the points at which the software analyses the data were below the initial starting baseline. As a result, the change in the refractive index, measured by the overlapping optical fringes was negative. And this is the reason why the software reports the output as errors. One potential method to overcome this problem would be to equilibrate the measured chips for an extended period of time to ensure the starting baseline is stable and no material is washed off the surface.

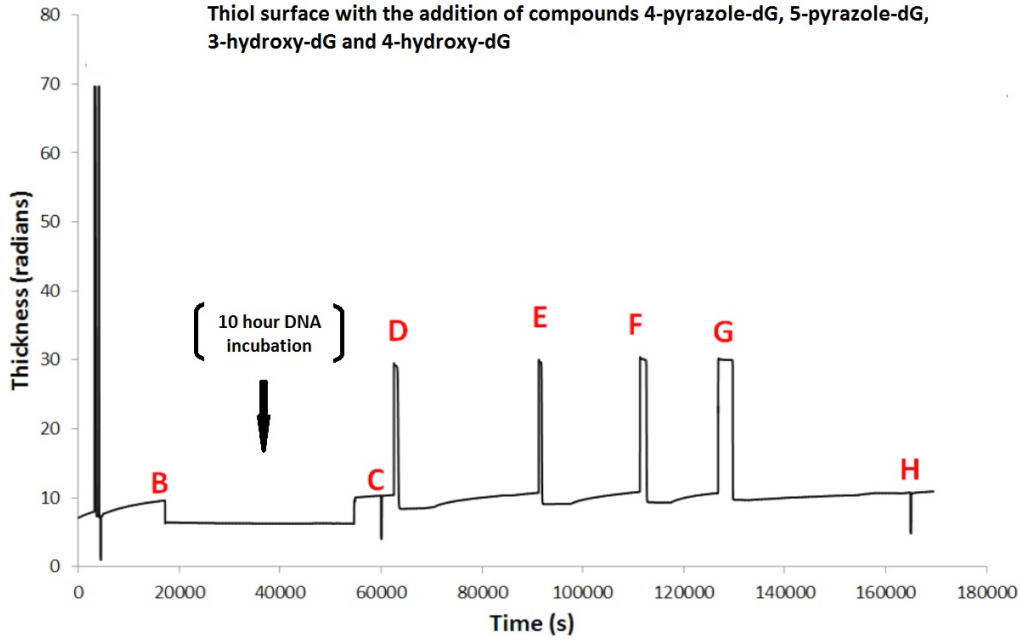


Figure 4.13: *A thiol-coated Anachip surface was calibrated using high purity ethanol and water as standards (A). 1µM 30-mer, thiolated, cy5-labelled DNA in H_2O was injected onto the Anachip surface and allowed to incubate for 10 hours (B). 5 mM potassium Iodide in H_2O was injected onto the Anachip surface (C). Four solutions of acetonitrile and water in a 1:1 mixture, containing 100 µM potassium Iodide and 100 µM of the compounds 3-hydroxy-dG (D), 4-pyrazole-dG (E), 5-pyrazole-dG (F) and 4-hydroxy-dG (G) were injected onto the surface of the Anachip. A solution of 5 mM sucrose in H_2O (H) was injected onto the Anachip surface.*

4.4.1.5 DPI Experiment 3

DPI experiment number 3 was performed on an amine-coated Anachip. Chip calibration was performed over a large period of time (as learnt from the previous experiment), 5 mM of Sulfo-SMCC linker dissolved in H_2O, was injected onto the amine-coated surface. The sulfo-SMCC appeared to bind successfully to the surface as evidenced by an increase in thickness to 1.0 nm, which is still within the measuring limits of the technique. This thickness actually corresponds to the length of the sulfo-SMCC molecule. Then, 1 µM of 30-mer thiolated telomeric DNA was injected onto the amine-coated surface at a flow rate of 5 mL/min. The DNA appeared to be bound to the activated surface, as indicated by the change in thickness 1.54 nm. This is only a very subtle change but it seems likely that Sulfo-SMCC must have been replaced in this

reaction and if the DNA is not oriented upright, the thickness of the single strand of nucleic acids (lying flat on the positively charged amino surface) is actually just below 2 nm. The immobilisation of the DNA was followed by injection of 100 µM solution of 4-pyrazole-dG with 100µM of potassium iodide in a 1:1 mixture of acetonitrile and water at 5 mL/min. A signal change was observed again, this time the thickness dropped from 1.5nm to 0.6 nm. Although this is again very subtle change, it may indicate a folding of the DNA into G-quadruplexes, induced by the addition of the 8ArdG. However, one would also expect that upon successful folding of telomeric DNA into G-quadruplexes, the mass and density of the material would also increase. However, the opposite was observed. The mass decreased from 0.83ng/mm^2 to 0.19ng/mm^2, and similarly the density decreased from 0.53g/cm^3 to 0.32g/cm^3. Although these results are contradicting, there is still a possibility that a small amount of DNA was non-specifically bound onto the Anachip surface. During the course of the incubation experiment with 8ArdG, some portion of the non-covalently bound DNA could have been removed. This would obviously contribute to the decrease in mass and surface density.

Another possible hypothesis to explain this phenomenon is an orientation change of the telomeric, G-quadruplex DNA as illustrated on figure 4.14.

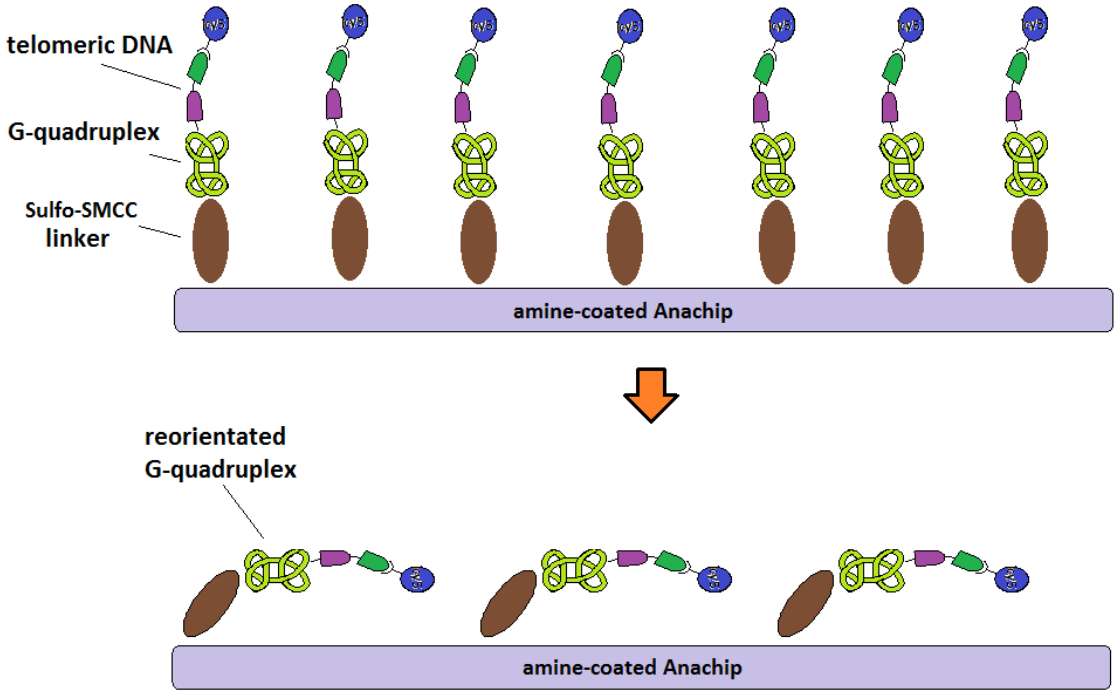


Figure 4.14: *Cartoon illustrating telomeric DNA reorientating to occupy space left after the departure of the Sulfo-SMCC linker from the Anachip surface.*

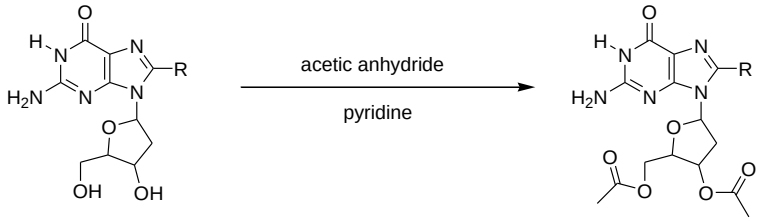
It is very difficult to prove or disprove any of the hypotheses provided. However, a final DPI experiment (#4) was performed on an amine-coated Anachip using a different 8ArdG. The incubation protocol was identical with the experiment 3 and the results were very similar. This

time, the thickness of the immobilised DNA was measured as 1.1 nm. An incubation of 100 µM solution of 5-pyrazole-dG with 100 µM of potassium iodide, in a 1:1 mixture of acetonitrile and water, injected at 5 mL/min lead to a decrease in thickness to 0.3 nm. When incubating 4-pyrazole-dG, the thickness change was of 0.9 nm. Upon injection of 5-pyrazole-dG, the thickness dropped by 0.8 nm. In both cases, the apparent decrease in thickness upon injection of the 8ArdG molecules was attributed to G-quadruplex formation. The expected changes in thickness were not as originally proposed and expected. Further experiments would need to be performed to validate these results and to confirm that the novel 8ArdG derivatives have the capacity to induce G-quadruplex formation on telomeric DNA molecules. In conclusion, the DPI technique is indeed a powerful method to resolve surface characteristics, with a resolution to 1 angstrom. It does have some limitations but as with many other techniques, it requires optimisation of many variables that are part of the measurement. Unfortunately, due to time constraints, this was not the case with many of the performed experiments. Too many times more than one variable was changed, which does not enable the deduction of meaningful conclusions from each test.

5.1 Future work

The following additional future experiments are presented as potential avenues for any subsequent research regarding the work previously presented in this thesis.

5.1.1 Future synthetic methods for the acetylation of 2`-deoxyguanosine derivatives



Scheme 5.1 *esterification of 5` and 3` OH of 8-aryl substituted 2`-deoxyguanosine derivatives*

The above reaction is a proposed methodology for the esterification of the 3`OH and 5`OH. The methodology was published by Anna Winqvist and Roger Strömberg[85]. The methodology uses pyridine as both the solvent and the base to deprotonate the 3` and 5` oh groups. The revised methodology also uses only 1x wash of NaCl saturated water. Reactions using this methodology have since been performed on 8-(3-Acetylphenyl)-2'-deoxyguanosine. Thin Layer Chromatographic analysis of the reaction mixture indicated the presence of the desired product.

5.1.2 Immobilization of thiolated Cy5 labelled DNA onto a gold surface

The immobilization of thiolated DNA was performed by depositing 20, 30 and 50 droplets of 500 picolitres of a solution of 1µM cy5-labelled thiolated DNA onto the surface of a gold-coated glass chip. The conclusion of the deposition experiments was that the DNA had not appeared to successfully bind to the gold surface.

An alternative method for the formation of the DNA monolayer over the gold surface could be achieved by adaptation of the technique used by the forester group[86]. The adapted methodology would involve the cutting of the gold coated 25mm x75mmx 1.1mm borosilicate glass chip into 10 25mm x7.5mmx 1.1mm smaller glass chips. As shown in figure 5.1., the smaller cut slides could then be inserted into a custom plastic slidetray (figure 5.1 **D**), purpose built using a makerbot replicator 2X 3D printer to the dimensions of the cut gold slide. The custom slide tray would allow for the analyses of DNA adhered to the surface of the cutdown slides with the PerkinElmer ScanArray Express instrument.

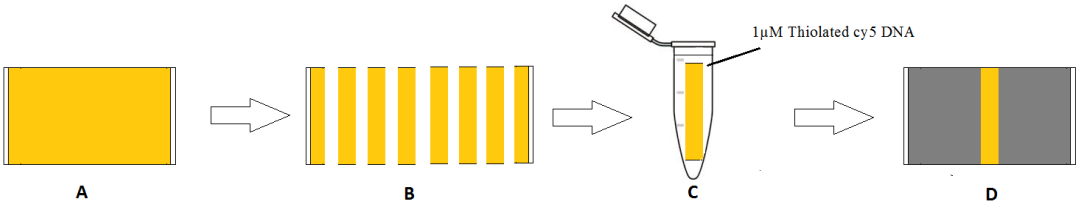


Figure 5.1 *A is the gold coated borosilicate glass chip. B is the glass chip cut into smaller sections of demensions 25mm x7.5mmx 1.1mm. C is a single glass chip piece submerged in a solution of 1µM thiolated cy5 DNA to allow monolayer assembly. D is the DNA coated gold slide piece in a 25mm x75mm x 1.2mm plastic holder ready for fluorescent analyses.*

Another possible approach for the adhesion of the thiolated DNA to the gold surface, could be to use techniques form a nature protocols publication[87]. In this publication the Plaxco lab developed a highly successful and reproducible method for the adhesion of single stranded DNA to a gold surface. This immobilised DNA is capable of detecting specific target analights, including small molecules, DNA and even cells in complex media such serum. The methodology is outlined in figure 5.1

The Plaxco group at the Chemistry and Biochemistry department of UCSB specialise in electrochemical DNA sensor development and are regarded as world leaders in their field. These DNA sensors are formed on gold. Signal generation is based upon conformational changes in single stranded probe DNA when it binds its complementary (target) strand. In the absence of target, the electrode bound single stranded redox tagged DNA is flexible and can easily fold in such a way that the redox tag strikes the electrode (gold) and generates a measurable current Figure 5.2. Upon target introduction, hybridisation between the probe and target occurs to form a double helix. This structural change results in a more rigid molecule and adversely effects the collision dynamics of the redox tag with the electrode surface and, ultimately, lowers the electron transfer rate between the two. This lowered rate is recorded as a drop in amperometric signal.

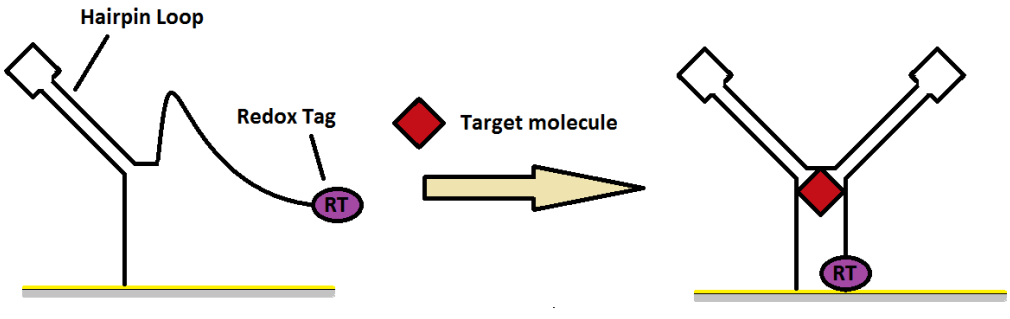


Figure 5.2: *Diagram of immobilised DNA, hairpin sensor on a gold surface, target molecule binds to the DNA, forming a hairpin conformation, which brings the redox tag (RT) in contact with the gold surface, completing the circuit and emitting a detectable electronic signal.*

This electrochemical sensor fabrication method has proven to be robust and could be modified to target deoxyguanosine molecules. Using a modified fluorescence-based approach, replacing the redox tag with a fluorophore, a single stranded fluorescently-labelled DNA's optical signal would be quenched by the gold surface in the absence of target and would exhibit fluorescence upon introduction once the target 2`-deoxyguanosine is present.

Outcomes and Conclusions

The first goal of the project was the synthesis of a library of novel compounds. A library of 4 compounds, 4-hydroxy-dG, 3-hydroxy-dG, 4-pyrazole-dG and 5-pyrazole-dG was successfully synthesised and fully characterised by IR, ^{1}H, ^{13}C, COSY, HMBC and HMQC NMR. A crystal X-ray structure of the compound 4-pyrazole-dG was obtained.

The second goal of the project was the synthesis of a covalently linked tetrad. A key step in this goal was the synthesis of the compound 4-ethynlylphenyl-dG, the purification of which proved problematic. Another key step was the synthesis of a linker functional group in the 5` position of deoxyguanosine. The synthetic approach attempted in this research was the functionalization of the 5` alcohol to an azide, for use in a Huisgen cycloaddition for coupling of deoxyguanosine subunits to form the CLGT. The substitution of the 5`OH proved problematic, despite repeated attempts. For future attempts at the synthesis of a CLGT an alternative strategy of coupling the subunits via a functional group linker at C^{8} position could be attempted as a potentially, more effective method.

The third goal of the project was the development of a novel screening method for the assessment of GQ formation, induced by 8ArdG analogues prepared in the first phase of the project. The methodology was based on the fluorescent quenching of Cy5-labelled telomeric DNA on a gold surface. The library of compounds synthesised in the first half of the project was used to test the new screening methodology. The analysis of the results of the experiments were inconclusive, one plausible hypothesis indicated that the method of depositing DNA resulted in the non-specific binding of the DNA in a parallel orientation to the gold surface with associated quenching of the fluorescent signal. The compounds were deposited onto the surface-bound DNA, in an attempt to induce GQ formation and draw the fluorophore away from the gold surface. However, as the results were inconclusive it is possible that the DNA was not successfully immobilised to the gold surface. The results showed that it was not possible to test the novel detection approach because of the inability to confirm the successful immobilisation of DNA molecules to the gold surface. The final goal of the project was analysis of the compounds for their ability to induce GQ formation in telomeric DNA by DPI. Again, the experimental results were not fully conclusive. However, there was some evidence

that the compounds 4-pyrazole-dG and 5-pyrazole-dG could induce the formation GQ in the presence of covalently surface bound DNA on an amine-coated Anachip. Overall, this project has potential to be developed to successful stage. It would require very careful planning, organisation of the work and significant effort to perform and optimise many reaction parameters.

Appendices

Appendix 1

Section 1: Infrared Spectroscopic details of Compounds 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG

The infrared spectrum of compounds 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG was measured on a Perkin Elmer spectrum 100 ATR FT-IR spectrometer. A small quantity of sample, approximately 3-5 mg was weighed out and paced on the quartz window. The diamond tip was firmly screwed down until the sample compound was immobilised. The IR The spectrum of the compound was measured between 600 to 4000 cm^{-1}, with a scan rate of 1 cm^{-1} per scan.

5-Pyrazole-dG

Peak (cm^{-1})	Absorbance	Functional group
3406 (broad)	81.9	-NH
3163 (broad)	58.9	-OH
1630 (sharp)	35.8	N-C=O
1594 (sharp)	29.5	Aromatic

4-Pyrazole-dG

Peak (cm^{-1})	Absorbance	Functional group
3403 (broad)	77.5	-NH
3194 (broad)	57.9	-OH
1649 (sharp)	42.1	-N-C=O
1601 (sharp)	29.1	Aromatic

4-Hydroxy-dG

Peak (cm^{-1})	Absorbance	Functional group
3406 (broad)	81.91	-NH
3163 (broad)	58.93	-OH
1630 (sharp)	35.78	N-C=O
1594 (sharp)	29.45	Aromatic

3-Hydroxy-dG

Peak (cm^{-1})	Absorbance	Functional group
3323 (broad)	73.5	-NH
3214 (broad)	73.0	-OH
1690 (sharp)	61.4	N-C=O
1614 (sharp)	41.7	Aromatic

Section 2: Infrared Spectroscopic display data of Compounds 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG and 4-hydroxy-dG

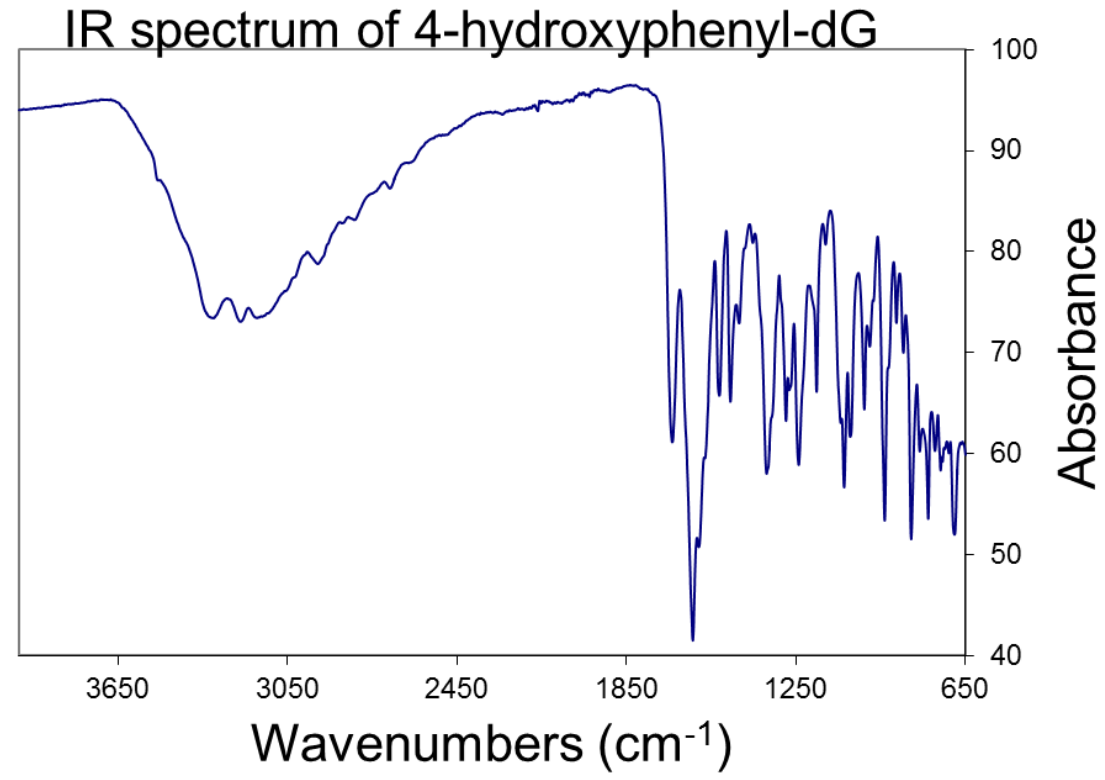


Figure A1.1: *IR spectrum of 4-hydroxyphenyl-dG*

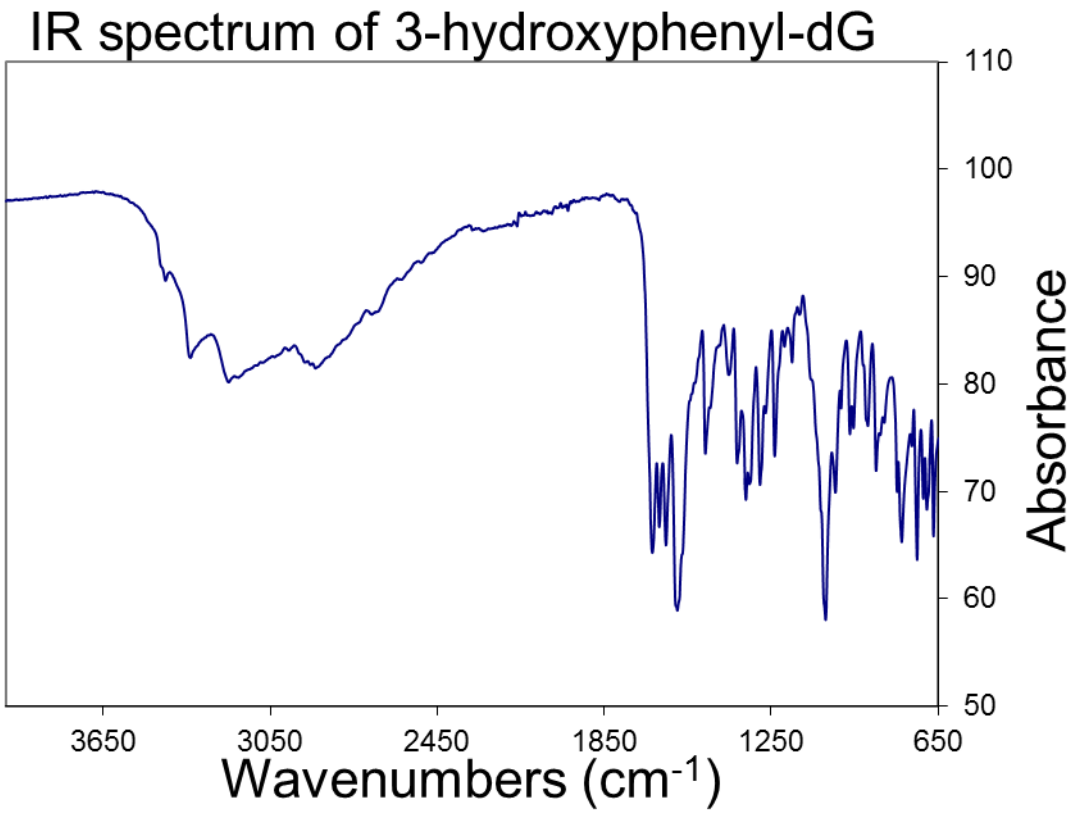


Figure A1.2: *IR spectrum of 3-hydroxyphenyl-dG*

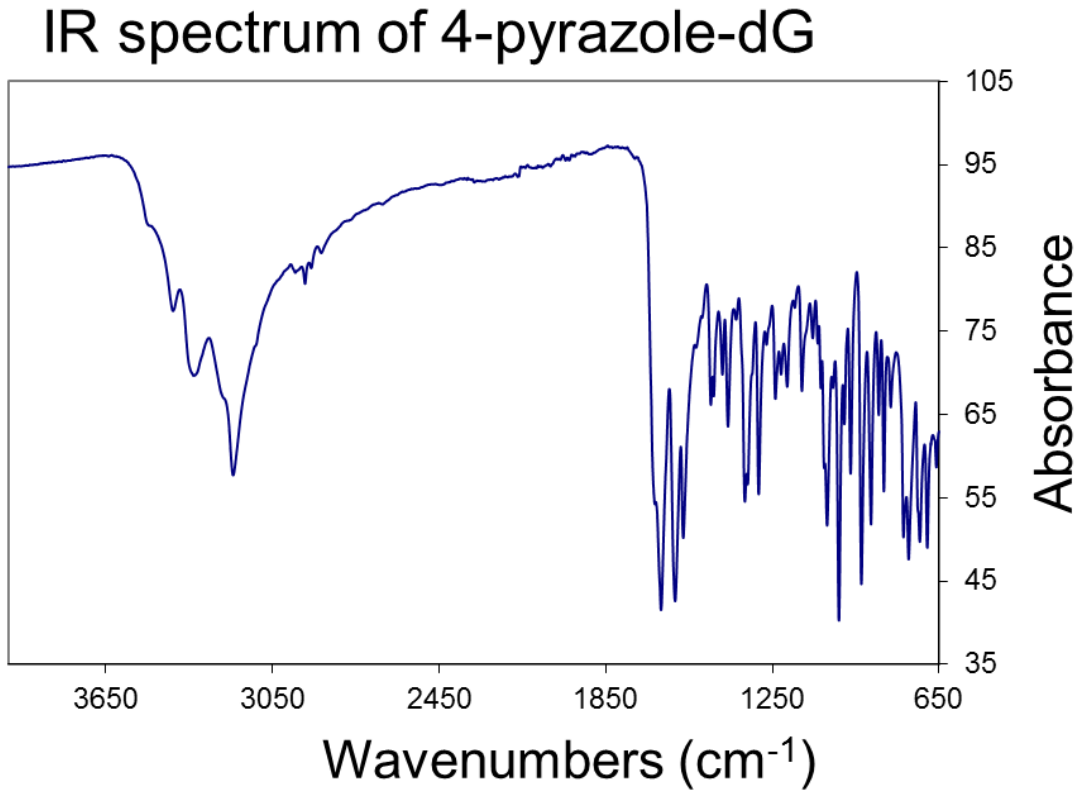


Figure A1.3: *IR spectrum of 4-pyrazole-dG*

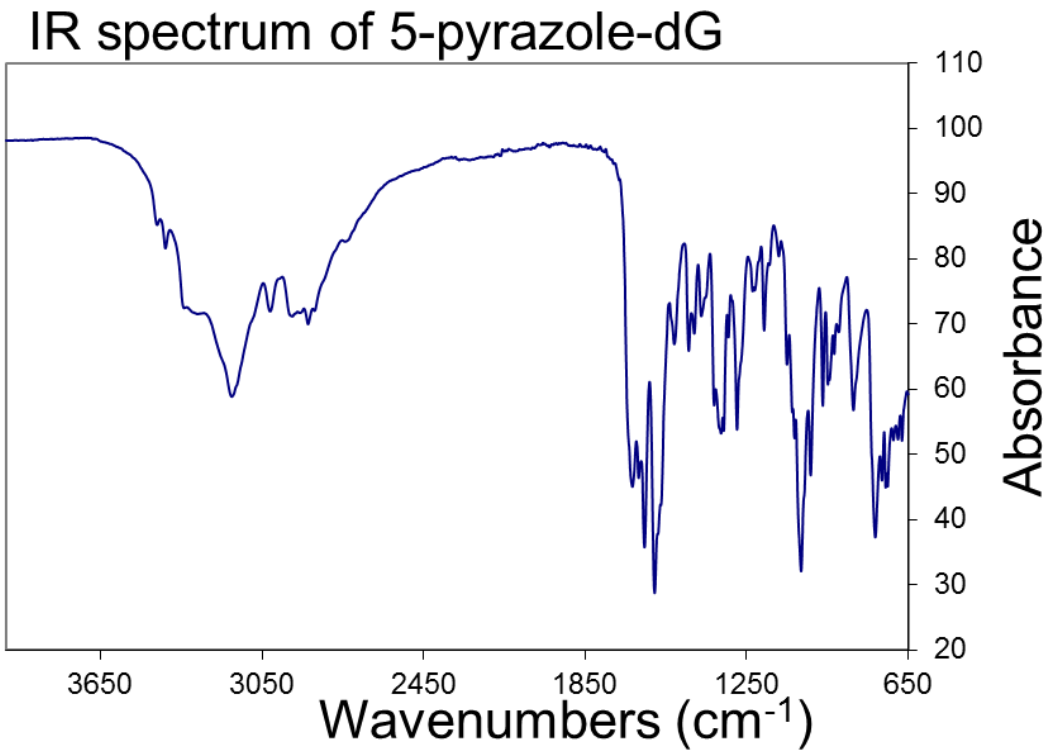


Figure A1.4: *IR spectrum of 5-pyrazole-dG*

Appendix 2

Section 1: Nuclear Magnetic Resonance Spectral Details of compounds 4-pyrazole-dG, 5-pyrazole-dG, 3-hydroxy-dG, 4-hydroxy-dG and 3-acetylphenyl-dG

4-pyrazole-dG

^{1}H NMR (400 MHz, DMSO)$_{\delta\text{-d6}}$ 2.102 (m, 1H, H2`), 3.137 (m, 1H, H2`), 3.604 (m, 2H, H5`), 3.819 (s, 1H, H4`), 4.440 (s, 1H, H3`), 5.190 (m, 2H, 5`OH/3`OH), 6.380 (s, 2H, NH_2), 6.761 (s, 1H, H``), 7.257 (s, 1H, H1`), 7.897 (s, 1H, H``), 10.862 (s, 1H, NH) 13.255 (s, 1H, NH``)

^{13}C NMR (125MHz, DMSO-d6): δ 37.327, 62.376, 71.510, 85.010, 87.834, 105.861, 117.271, 129.628, 140.959, 143.107, 151.757, 152.899, 156.646

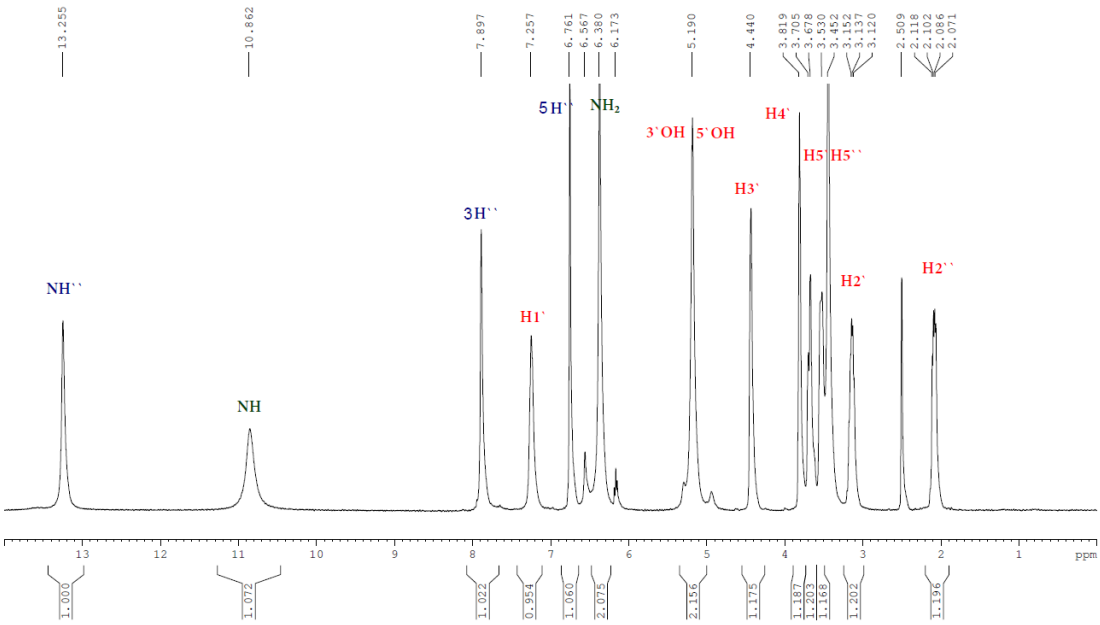


Figure A2.1: *^{1}H Nuclear Magnetic Resonance spectrum of 4-pyrazole-dG*

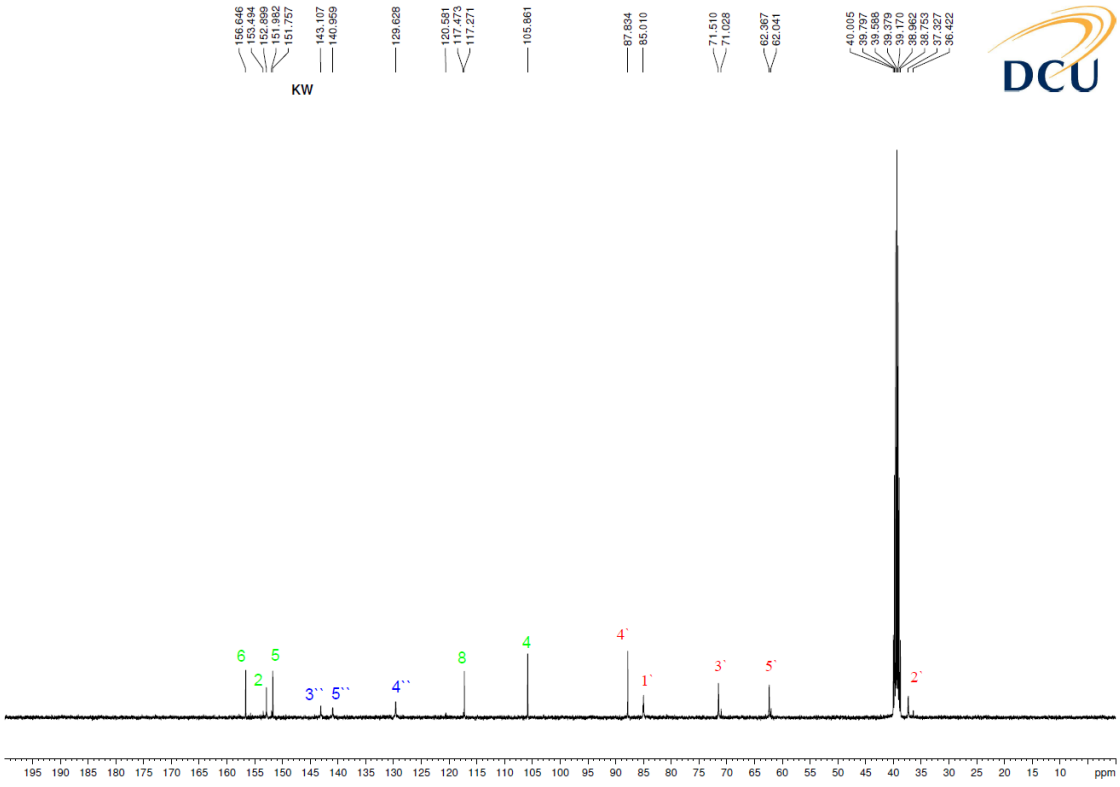


Figure A2.2: *^{13}C Nuclear Magnetic Resonance spectrum of 4-pyrazole-dG*

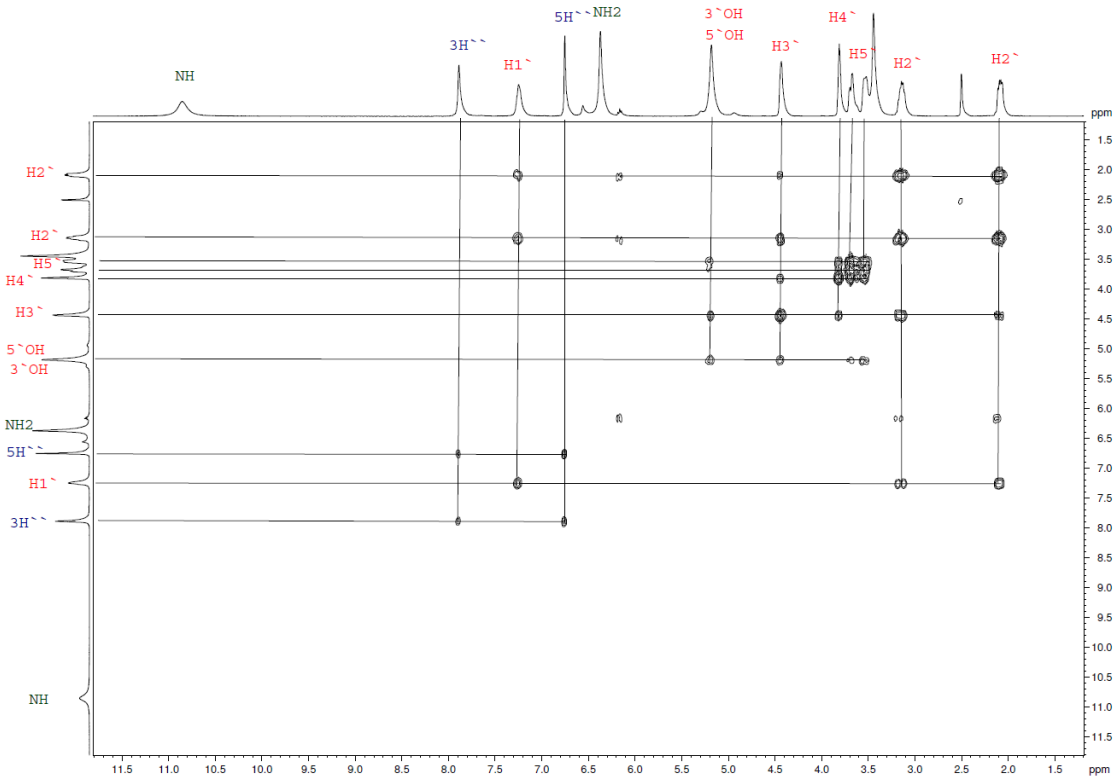


Figure A2.3: *Two dimensional Nuclear Magnetic Resonance correlation spectroscopy of 4-pyrazole-dG*

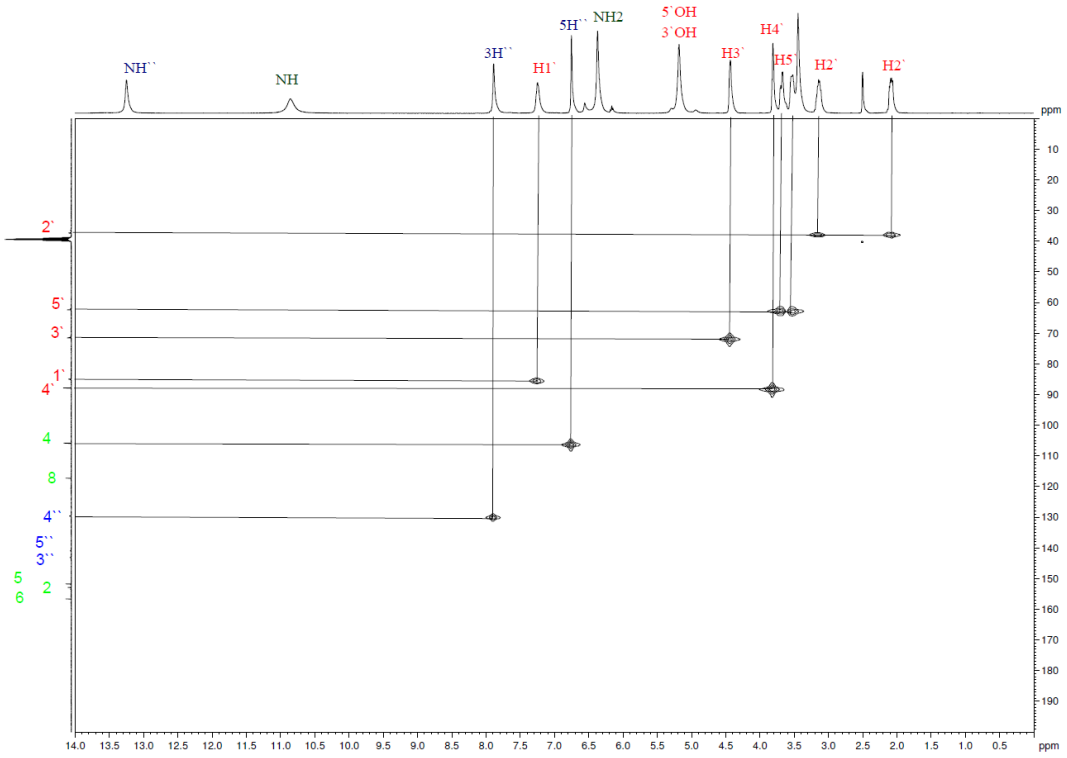


Figure A2.1: *Two dimensional Nuclear Magnetic Resonance heteronuclear multiple bond correlation spectroscopy of 4-pyrazole-dG*

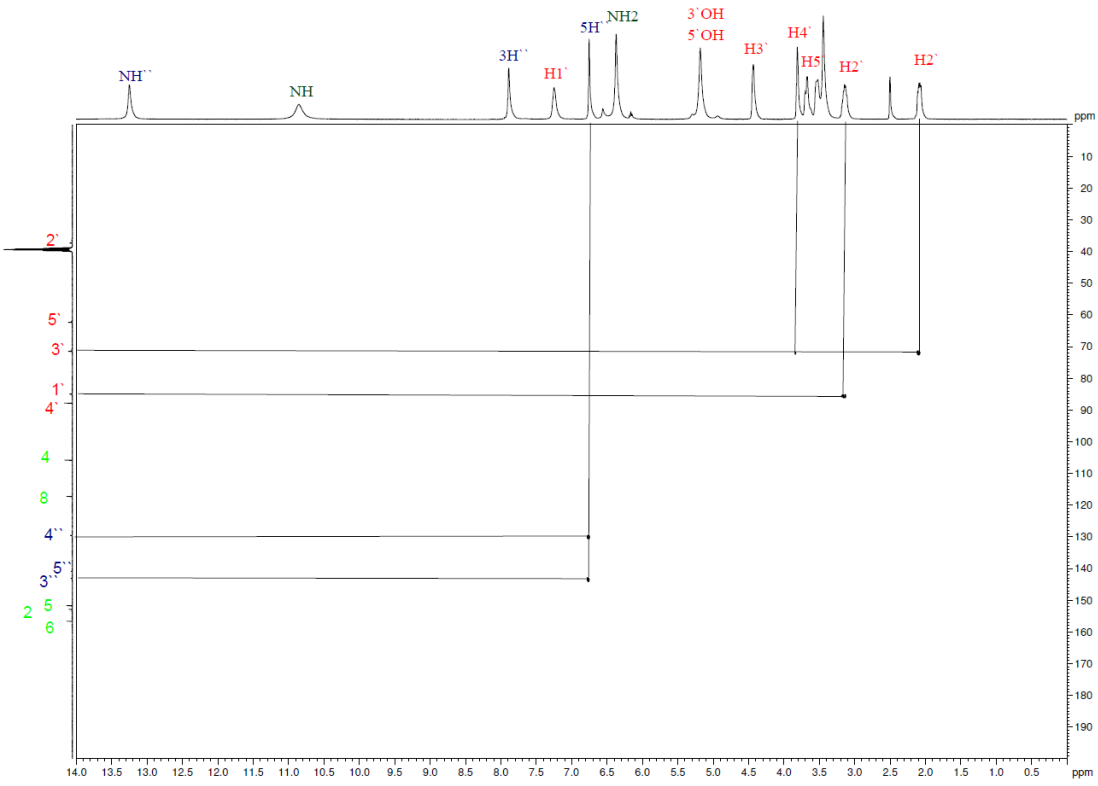


Figure A2.2: *Two dimensional Nuclear Magnetic Resonance heteronuclear single quantum coherence spectroscopy on 4-pyrazole-dG*

5-pyrazole-dG

^{1}H NMR (400 MHz, DMSO)$_{\delta\text{-d6}}$ 2.069 (m, H1, H2`), 3.221 (m, 1H, H2`), 3.583 (m, 2H, H5`), 3.811 (t, J=4 Hz, H4`), 4,378 (s, 1H, H3`), 5.057 (s, 1H, 5`OH), 5.236 (s, 1H, 3`OH), 6.159 (t, J=7.2 Hz, 1H, H1`), 6.409 (s, 2H, NH_2), 7.939 (s, 1H, 5H``), 8.078 (s, 1H, 3H``)

^{13}C NMR (125MHz, DMSO-d6): δ 36.408, 61.913, 70.994, 84.152, 87.643, 110.935, 116.764, 141.355, 151.591, 152.739, 156.527

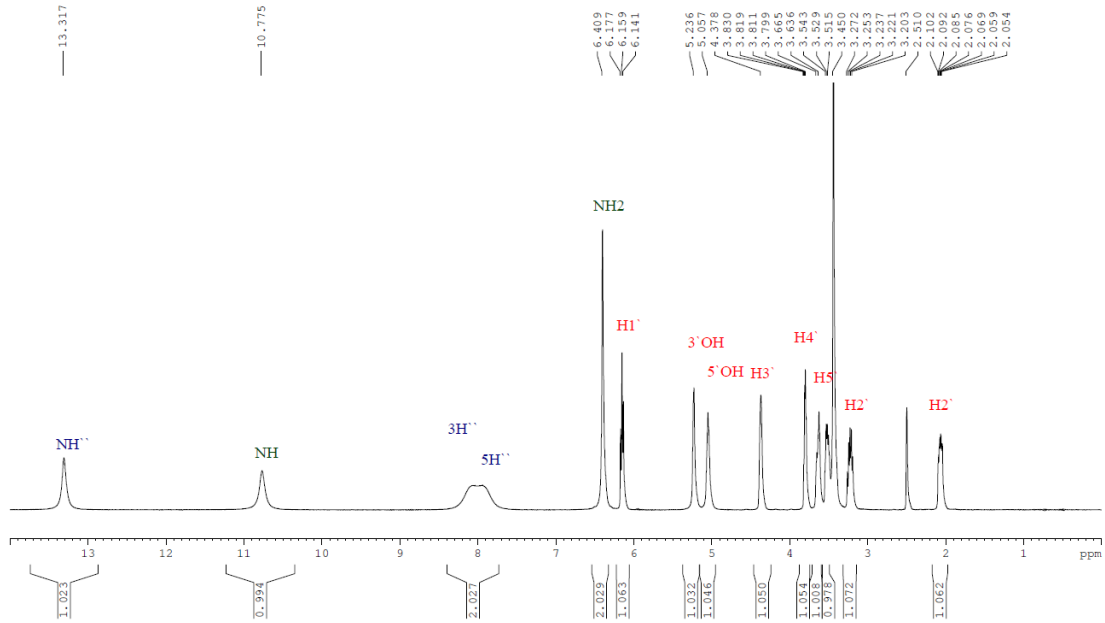


Figure A2.3: *^{1}H Nuclear Magnetic Resonance spectrum of 5-pyrazole-dG*

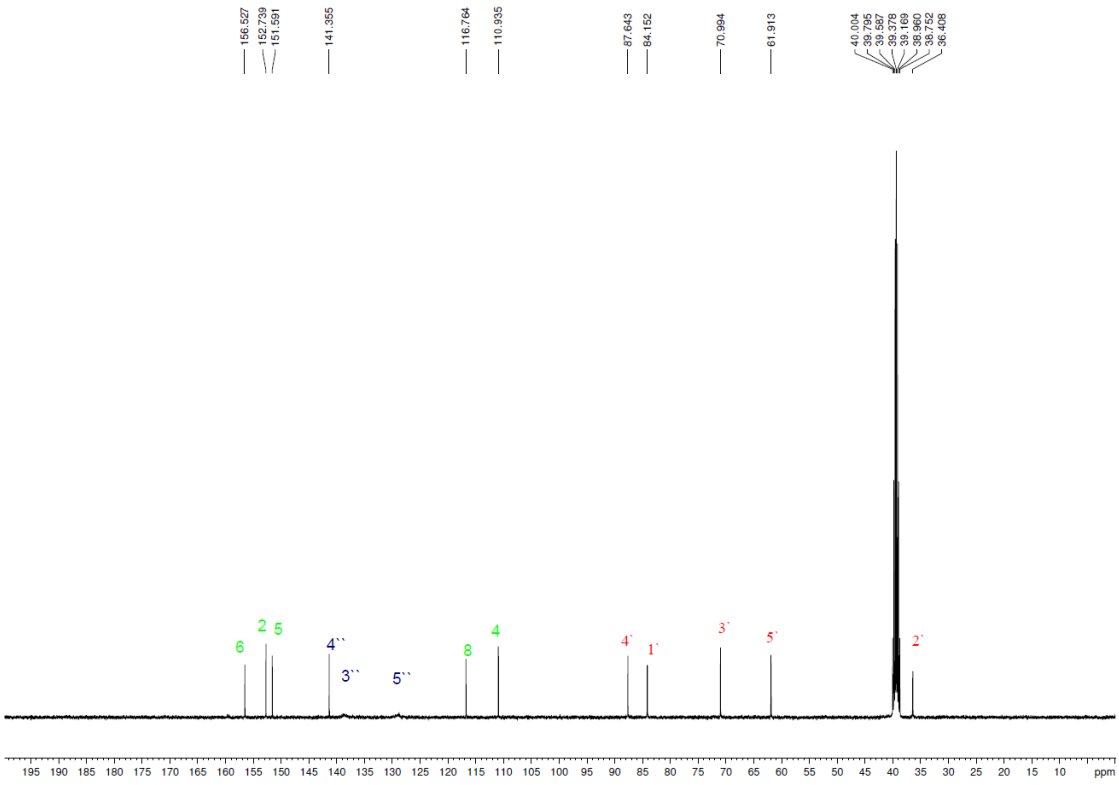


Figure A2.7: *^{13}C Nuclear Magnetic Resonance spectrum of 5-pyrazole-dG*

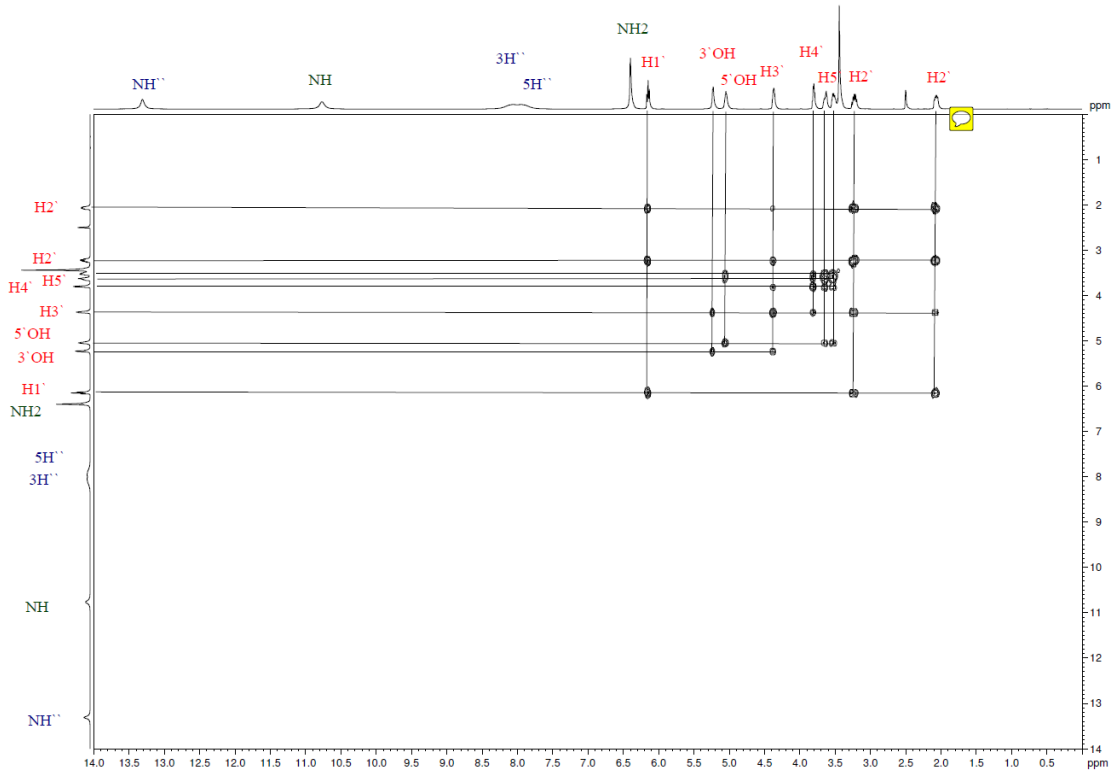


Figure A2.8: *Two dimensional Nuclear Magnetic Resonance correlation spectroscopy of 5-pyrazole-dG*

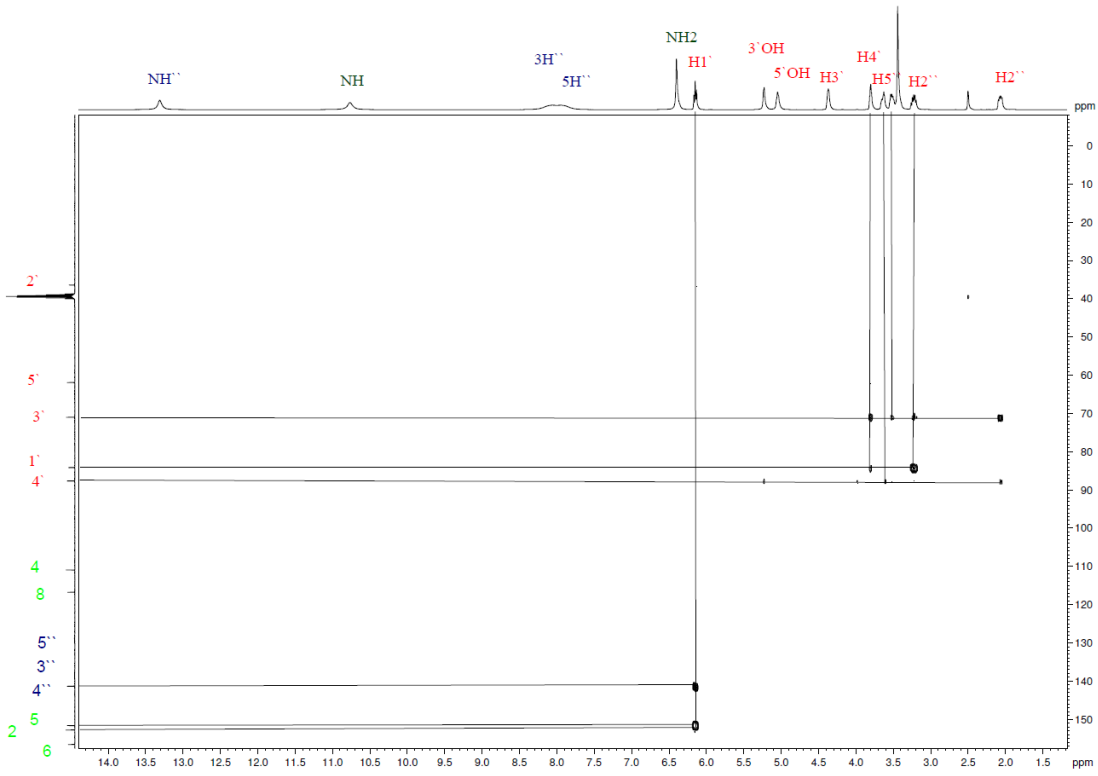


Figure A2.9: *Two dimensional Nuclear Magnetic Resonance heteronuclear multiple bond correlation spectroscopy of 5-pyrazole-dG*

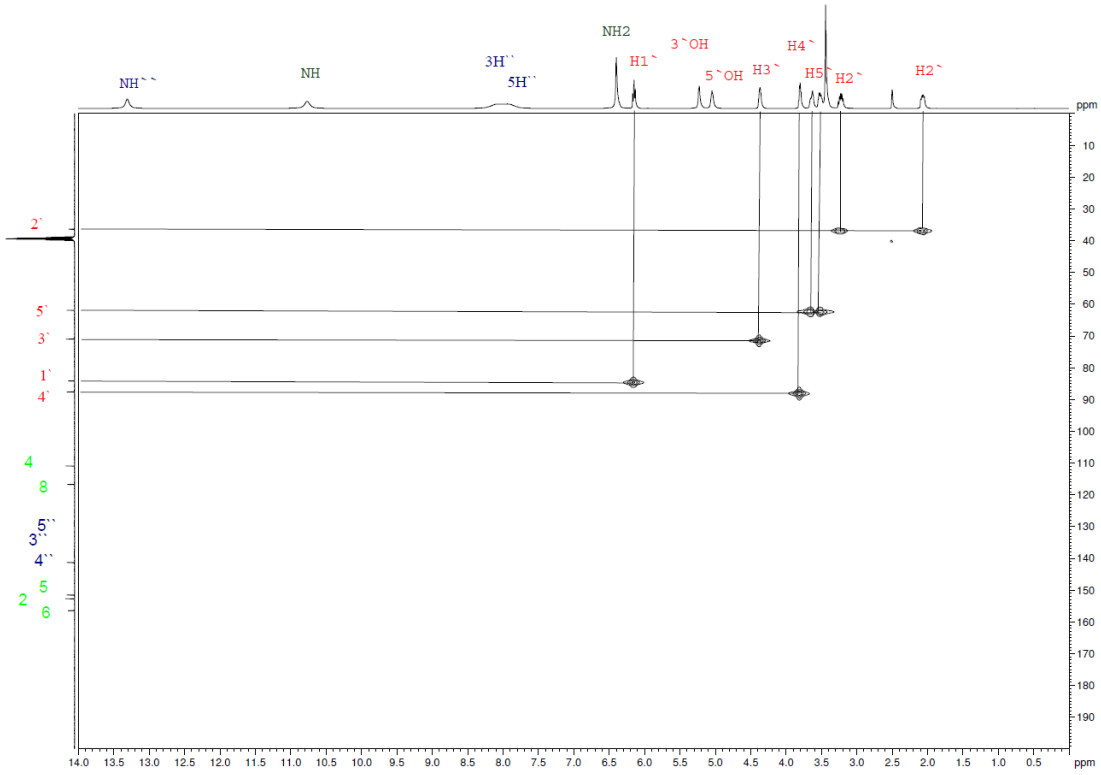


Figure A2.10*: Two dimensional Nuclear Magnetic Resonance heteronuclear single quantum coherence spectroscopy of 5-pyrazole-dG*

4-hydroxy-dG

^{1}H NMR (400 MHz, DMSO)$_{\delta\text{-}d6}$ 2.011 (m,H1, H2`), 3.160 (m, 1H, H2`) 3.603 (m, 2H, H5`), 3.791 (s, 1H, H4`), 4.340 (s, 1H, H3`), 5.062 (s, 1H, 5`OH), 5.192 (s, 1H, 3`OH), 6.052 (t, J=7.20 Hz, 1H, H1`), 6.399 (s, 2H, NH_2), 6.897 (d, J=8.4 Hz, 2H, H2`` H6``), 7.454 (d, J=8.4 Hz, H3`` H5``), 9.974 (s, 1H, NH), 10.750 (s, 1H, OH``)

^{13}C NMR (125MHz, DMSO-d6): δ 37.606, 60.971, 70.057, 86.623, 114.198, 115.661, 119.718, 129.499, 146.377, 150.619, 151.660, 157.392

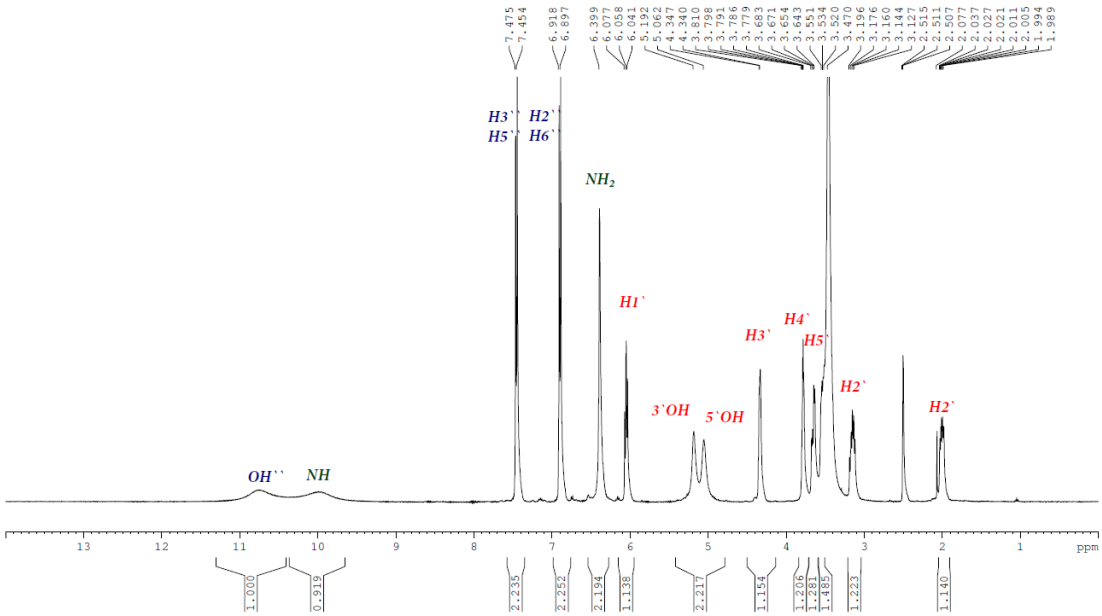


Figure A2.11: *^{1}H Nuclear Magnetic Resonance spectrum of 4-hydroxy-dG*

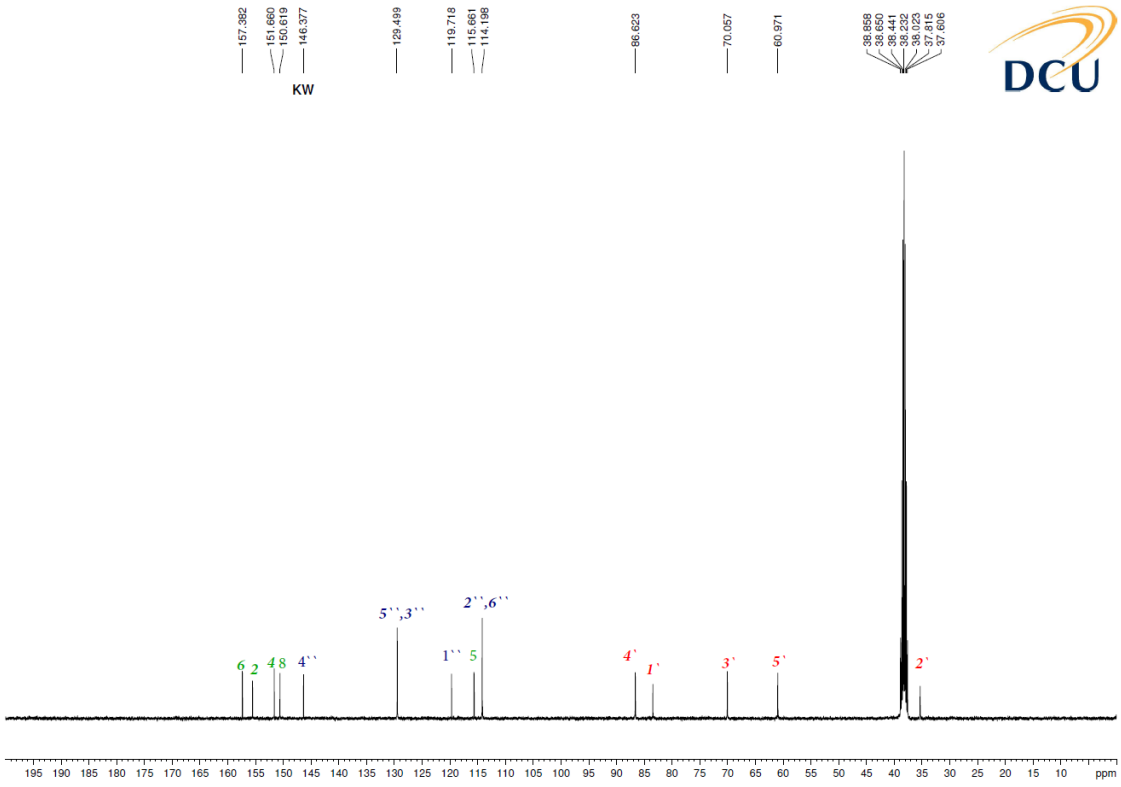


Figure A2.12: *^{13}C Nuclear Magnetic Resonance spectrum of 4-hydroxy-dG*

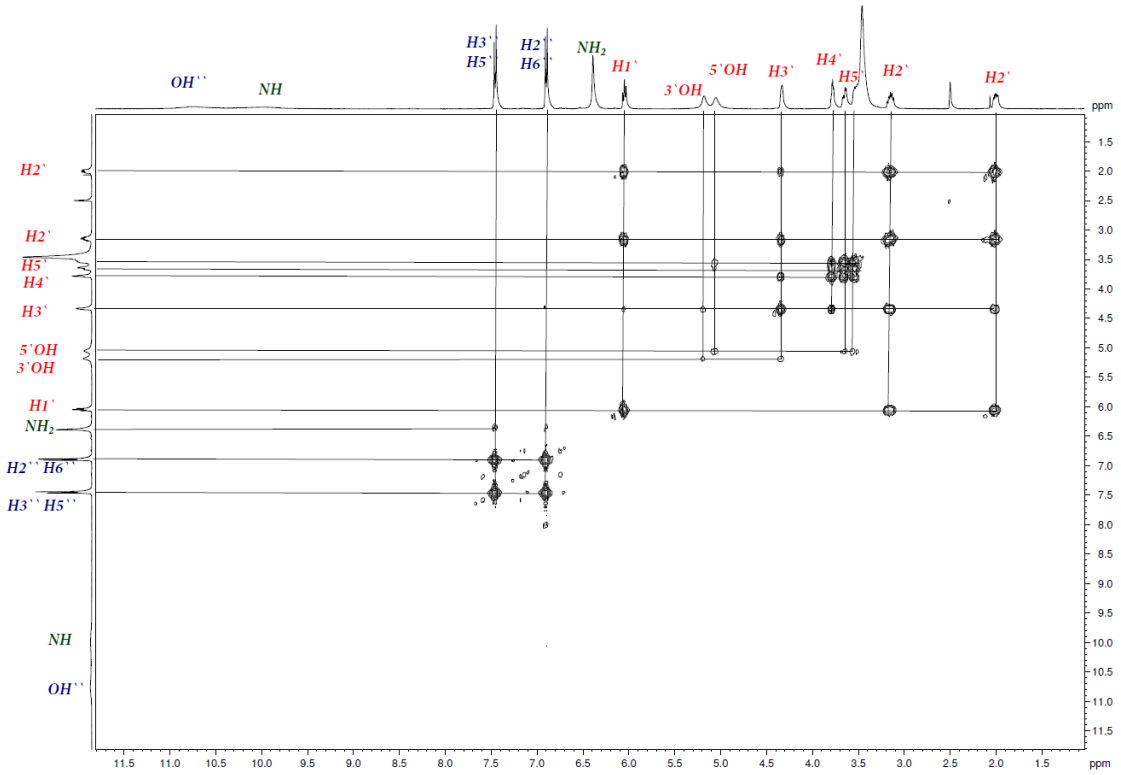


Figure A2.13: *Two dimensional Nuclear Magnetic Resonance correlation spectroscopy of 4-hydroxy-dG*

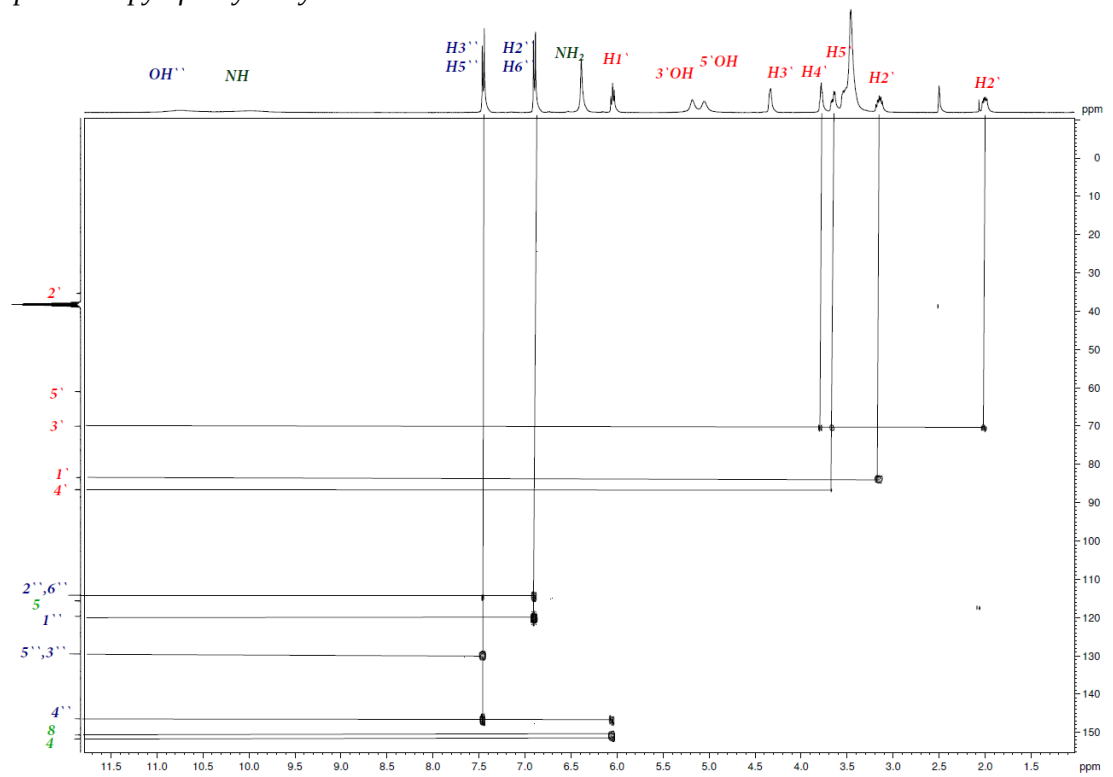


Figure A2.14: *Two dimensional Nuclear Magnetic Resonance heteronuclear multiple bond correlation spectroscopy of 4-hydroxy-dG*

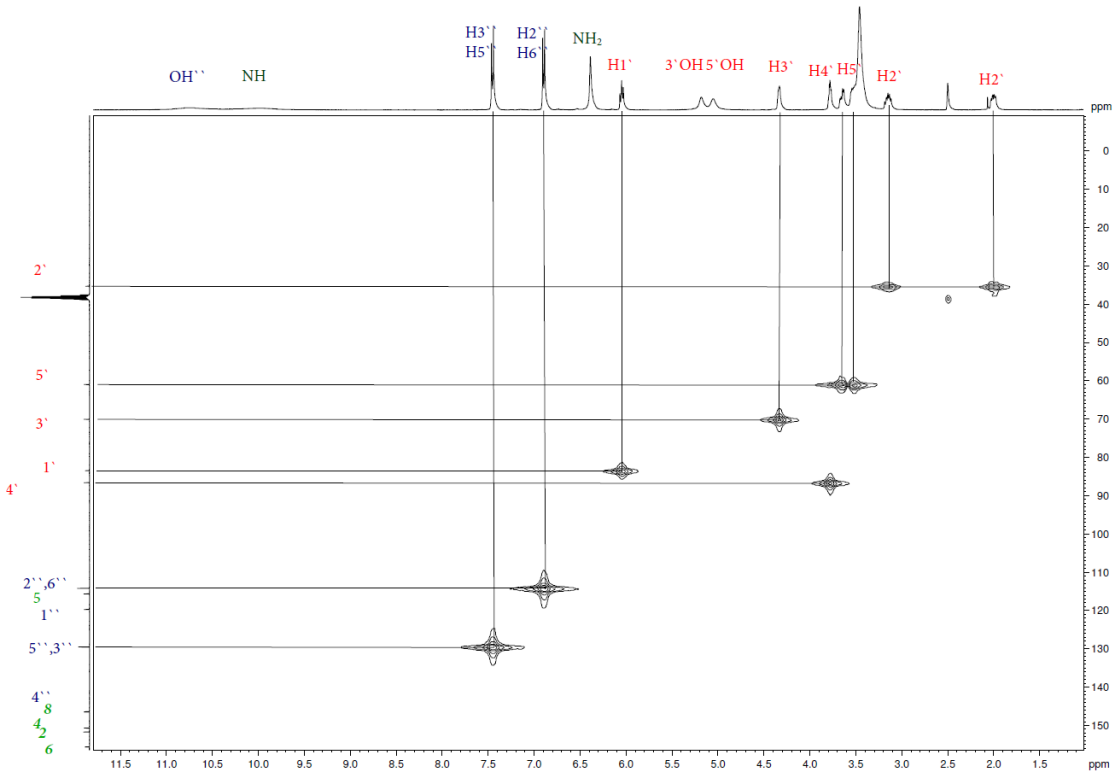


Figure A2.15: *Two dimensional Nuclear Magnetic Resonance heteronuclear single quantum coherence spectroscopy of 4-hydroxy-dG*

3-hydroxy-dG

^{1}H NMR(400 MHz, DMSO)$_{\delta\text{-}d6}$ 2.032 (m, H1, H2`), 3.209 (m, 1H, H2`), 3.627 (m, 2H, H5`), 3.811 (s, 1H, H4`), 4.370 (s, 1H, H3`) 5.044 (s, 1H, 5`OH), 5.201 (s, 1H, 3`OH), 6.106 (t J=7.544 Hz, H1`), 6.433 (s, 2H, NH_2), 6.926 (d, J=7.00 Hz, 1H, 6H``), 7.032 (s, 1H, 2H``), 7.051, (s, 1H, 4H``), 7.341 (t, J=7.8 Hz, 1H, 5H``), 9.826 (s, 1H, NH), 10.820 (s, 1H, OH``)

^{13}C NMR (125MHz, DMSO-d6): δ 36.475, 62.169, 71.288, 84.712, 87.909, 115.867, 116.437, 117.018, 119.731, 129.677, 131.312, 147.146, 151.871, 152.298, 156.298, 157.342

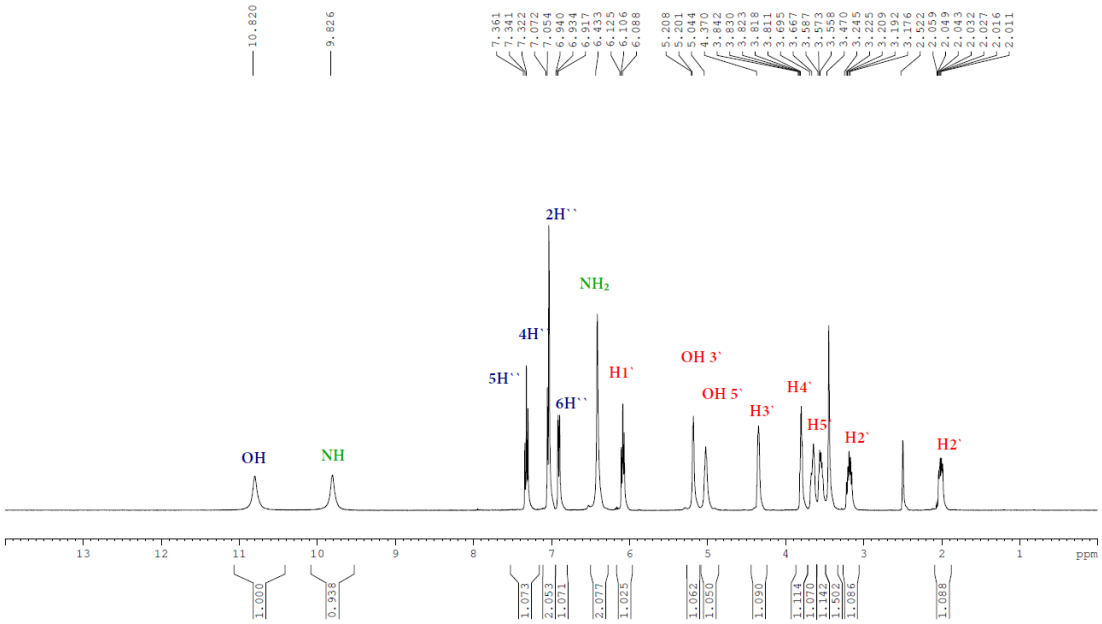


Figure A2.16: *^{1}H Nuclear Magnetic Resonance spectrum of 3-hydroxy-dG*

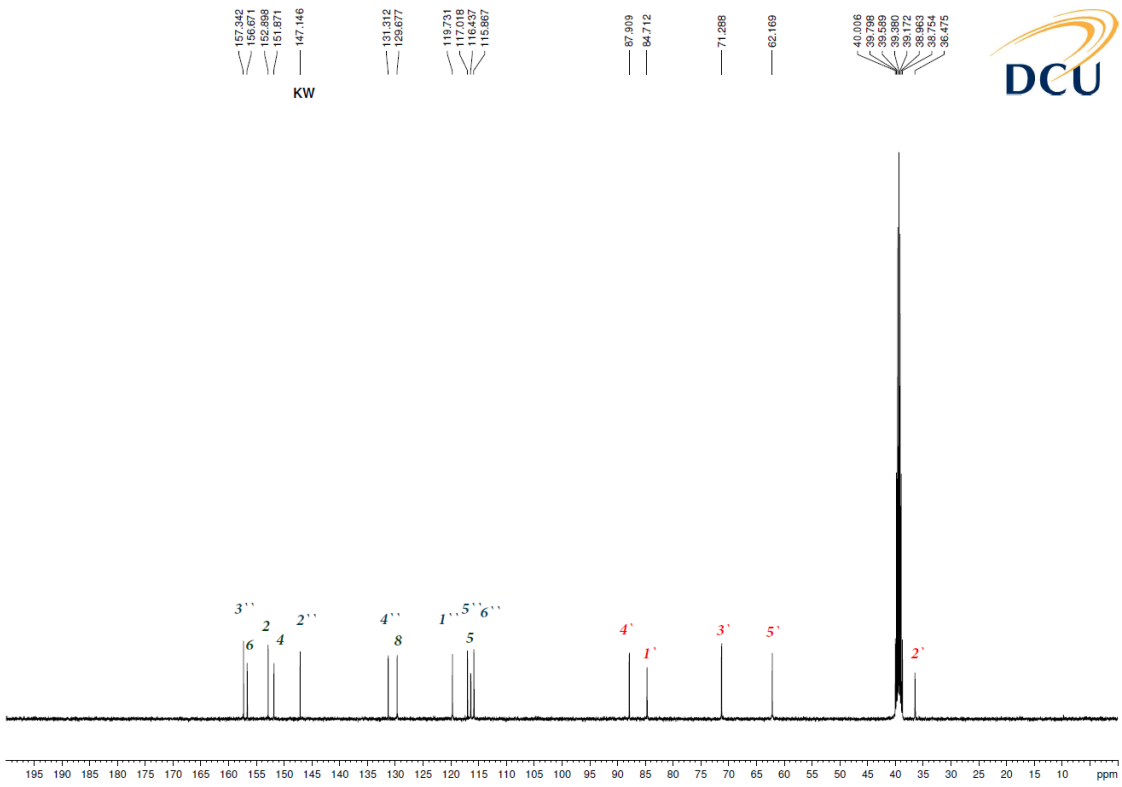


Figure A2.17: *^{13}C Nuclear Magnetic Resonance spectrum of 3-hydroxy-dG*

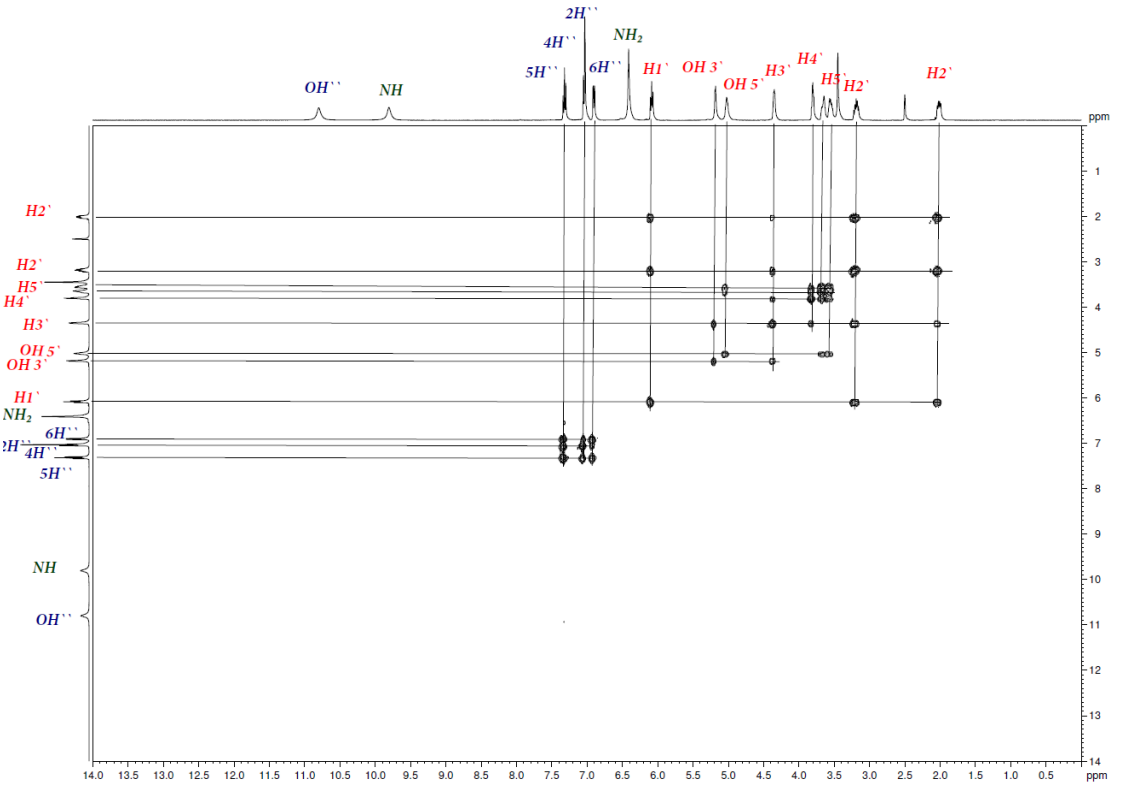


Figure A2.18: *Two dimensional Nuclear Magnetic Resonance correlation spectroscopy of 3-hydroxy-dG*

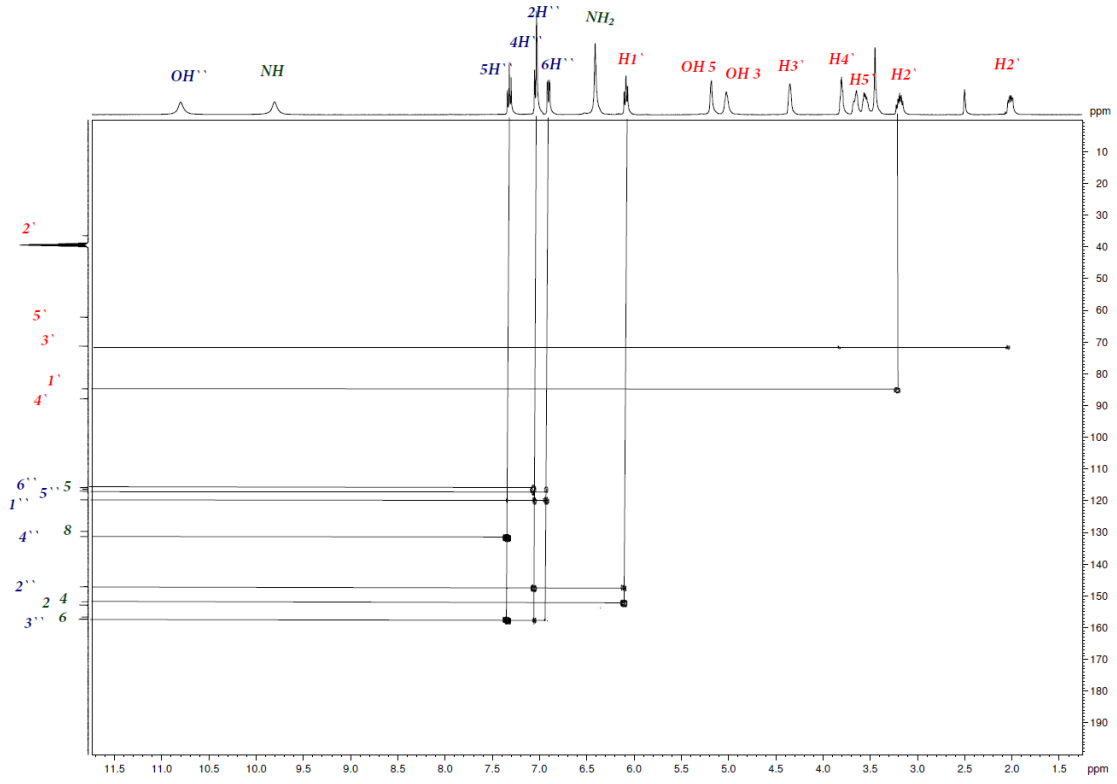


Figure A2.19: *Two dimensional Nuclear Magnetic Resonance heteronuclear multiple bond correlation spectroscopy of 3-hydroxy-dG*

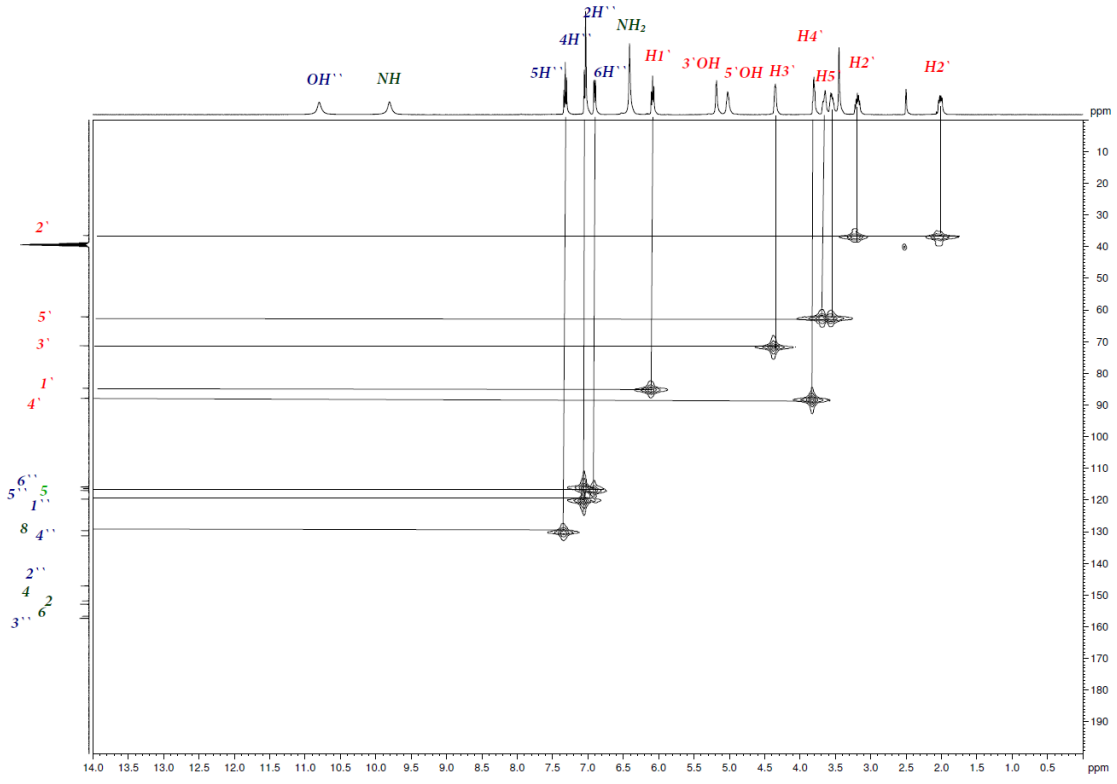


Figure A2.20: *Two dimensional Nuclear Magnetic Resonance heteronuclear single quantum coherence spectroscopy of 3-hydroxy-dG*

8-(3-Acetylphenyl)-2'-deoxyguanosine, (3-Ac-dG)

^{1}H-NMR (400 MHz, DMSO-d_6) δ 10.79 (s, 1H), 8.19 (s, 1H), 8.06 (d, J= 7.8 Hz, 1H), 7.90 (d, J=7.8 Hz, 1H), 7.68 (t, J=7.7 Hz, 1H), 6.45 (s, 2H), 6.04 (t, J=7.5 Hz, 1H), 5.13 (d, J=4.3 Hz, 1H), 4.93 (t, J=5.8 Hz, 1H), 4.30 (d, J=2.7 Hz, 1H), 3.77 (q, J=2.9, 5.3 Hz, 1H), 3.61 (m, 1H), 3.52 (m, 1H), 3.32 (s, 3H), 3.17 (m, 1H), 2.02 (ddd, J= 2.7, 6.5, 8.2 Hz, 1H).

^{13}C-NMR (100 MHz, DMSO-d_6) δ 197.97, 157.14, 153.60, 152.57, 146.68, 143.58, 137.51, 133.89, 131.25, 129.57, 123.62, 117.71, 88.35, 85.02, 71.58, 62.51, 39.49, 27.31.

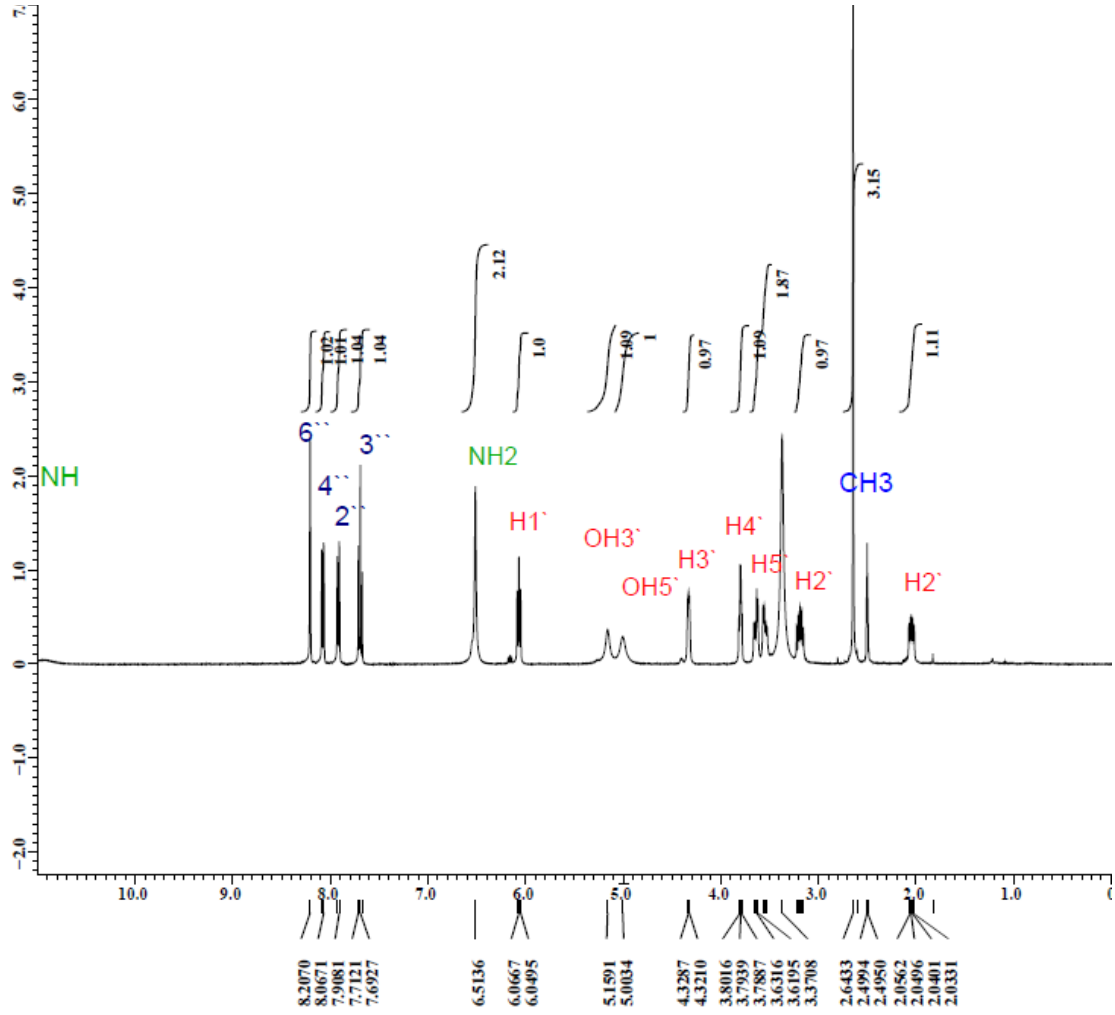


Figure A2.21: *^{1}H Nuclear Magnetic Resonance spectrum of 3-Ac-dG*

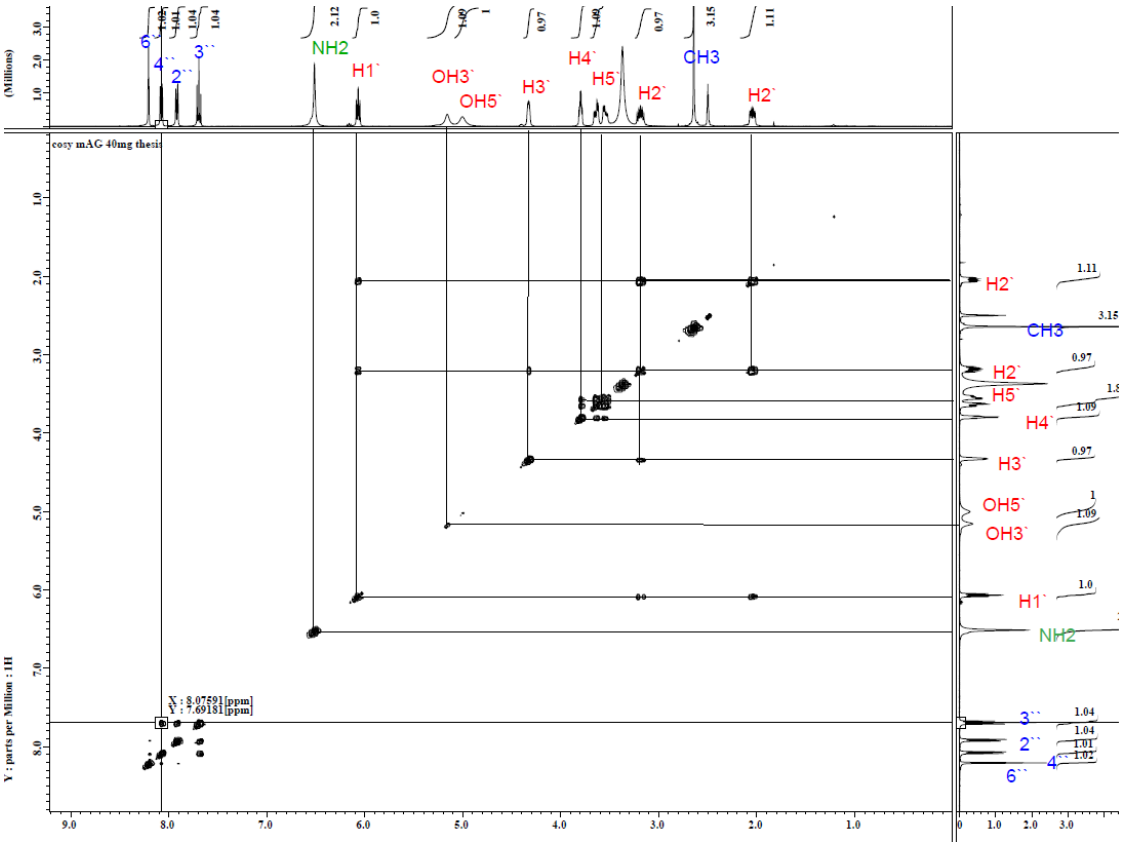


Figure A2.22: *Two dimensional Nuclear Magnetic Resonance correlation spectroscopy of 3-Ac-dG*

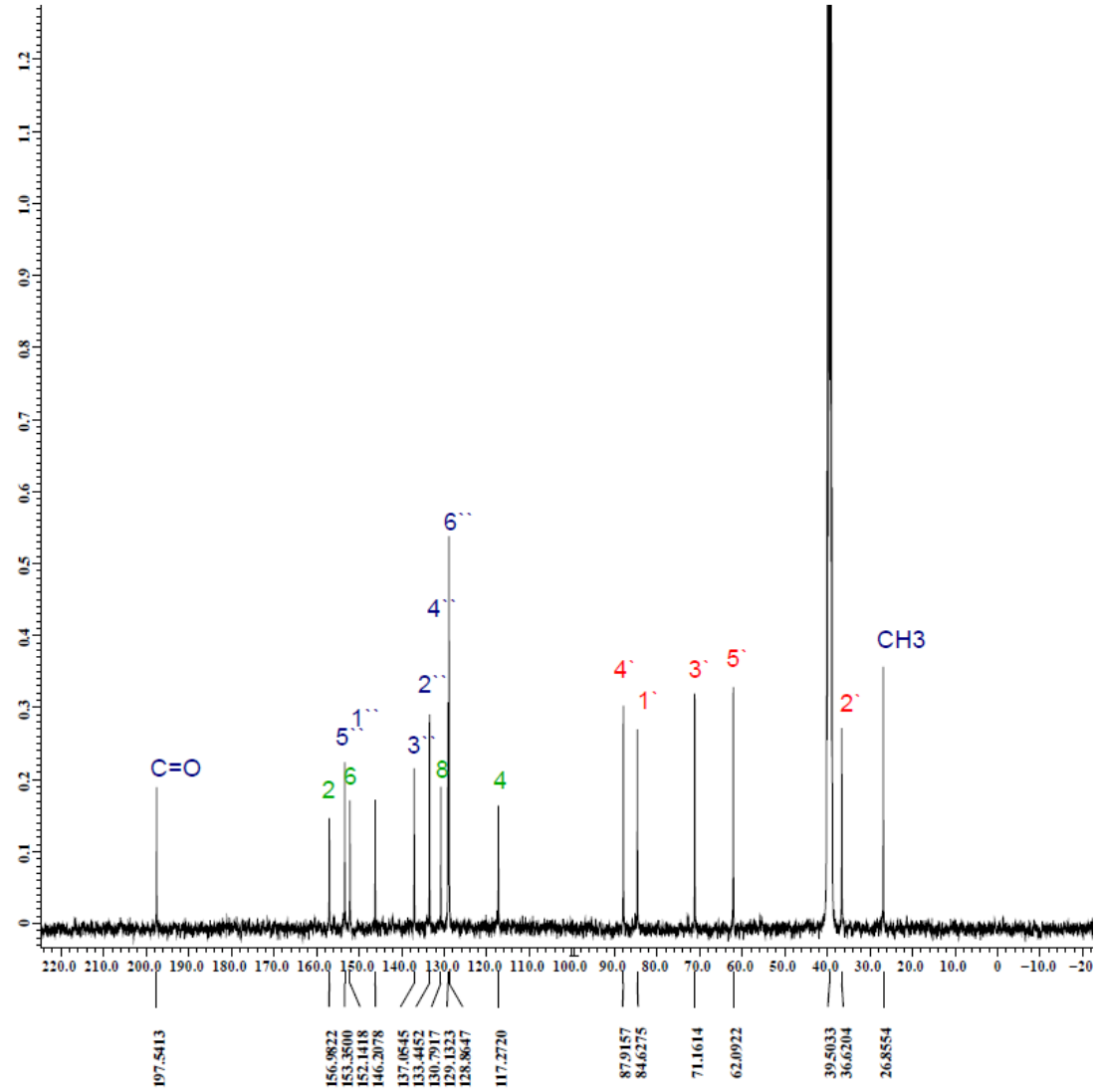


Figure A2.23: *Two dimensional Nuclear Magnetic Resonance heteronuclear multiple bond correlation spectroscopy of 3-Ac-dG*

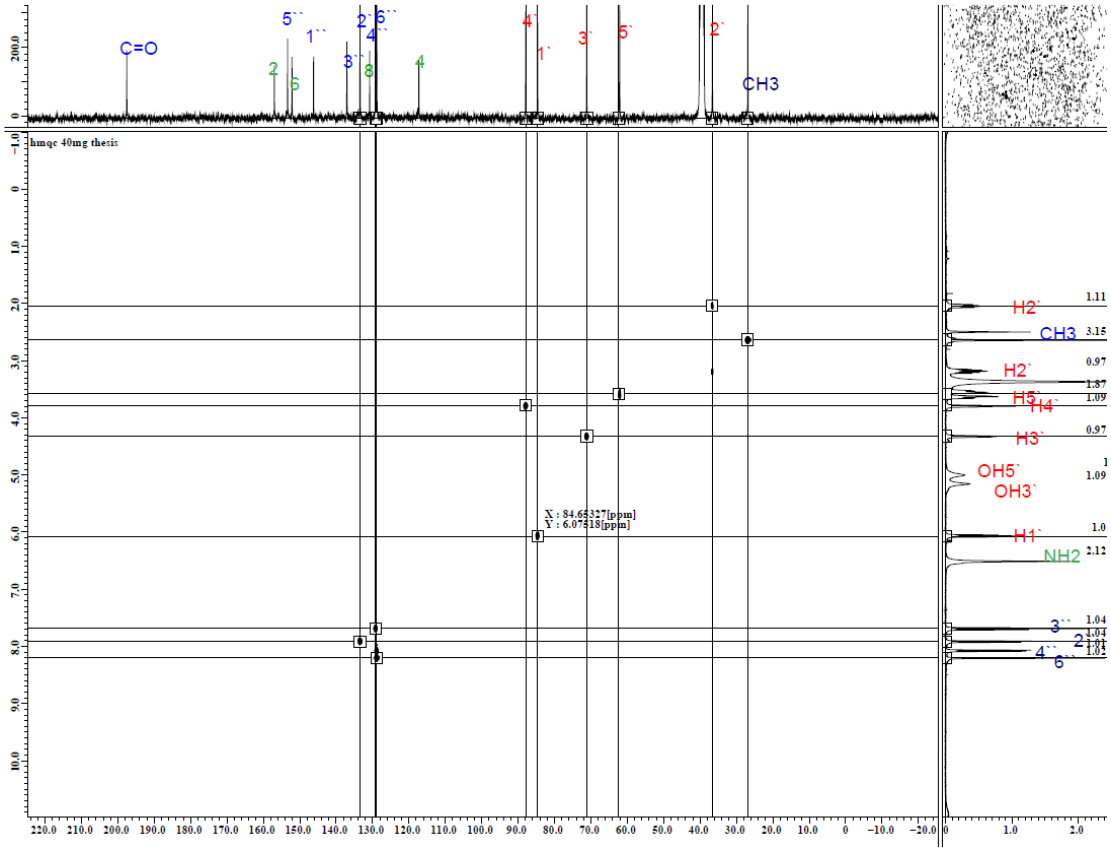


Figure A2.24: *Two dimensional Nuclear Magnetic Resonance Heteronuclear single quantum coherence spectroscopy of 3-Ac-dG*

Section 2: Nuclear Magnetic Resonance Spectral Details of the Appell reaction (3) and Mesylation and Tosylation reactions (4)

Reaction 3 Appell

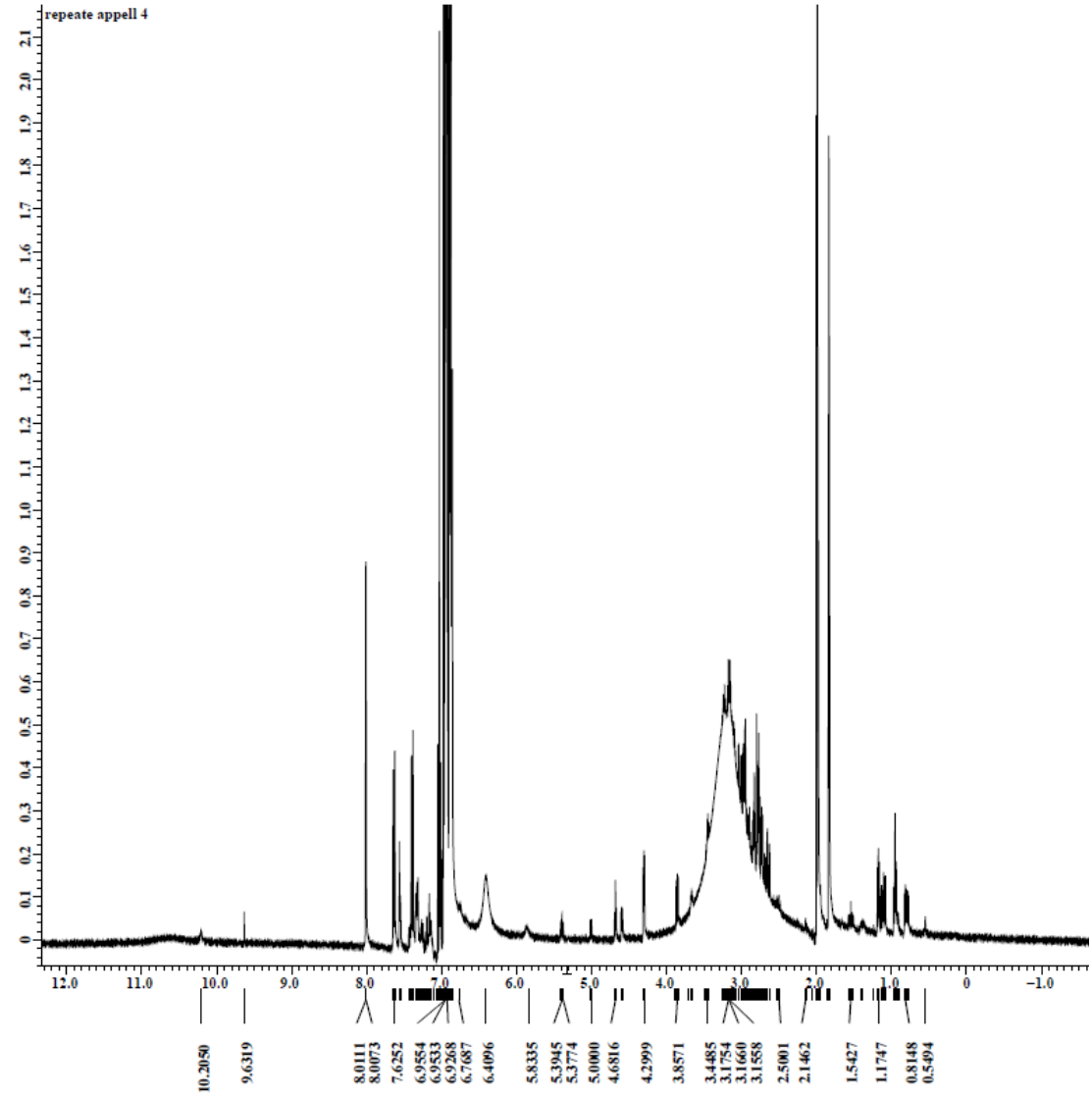


Figure A2.24: *^{1}H Nuclear Magnetic Resonance spectra for Appell crude 1.2 equivalents*

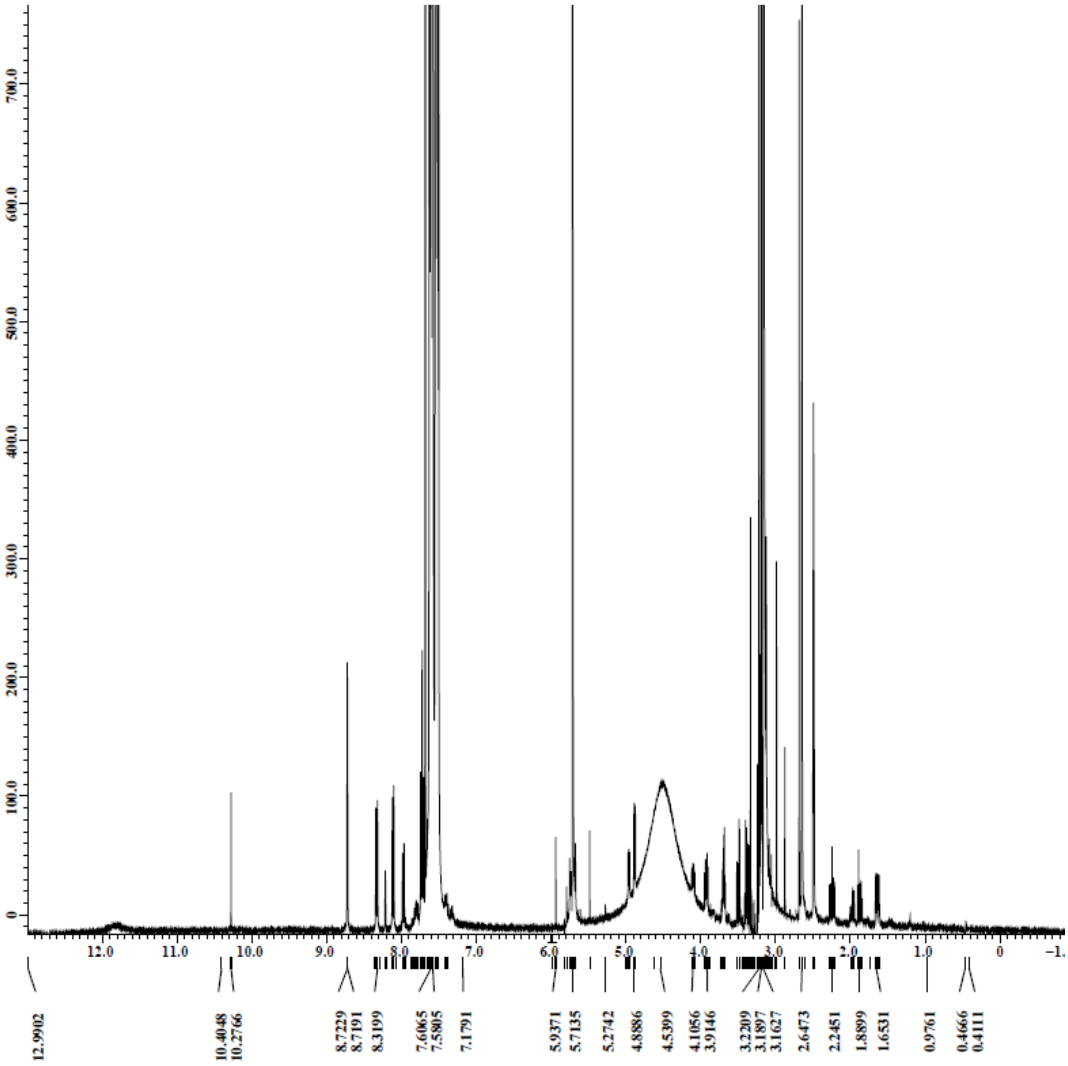

700.0
600.0
500.0
400.0
300.0
200.0
100.0
0
12.0
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0
12.9902
10.4048
10.2766
8.7229
8.7191
8.3199
7.6065
7.5805
7.1791
5.9371
5.7135
5.2742
4.8886
4.5399
4.1056
3.9146
3.2209
3.1897
3.1627
2.6473
2.2451
1.8899
1.6531
0.9761
0.4666
0.4111

Figure A2.25: *^{1}H Nuclear Magnetic Resonance spectra for Appell crude 1.2 equivalents*

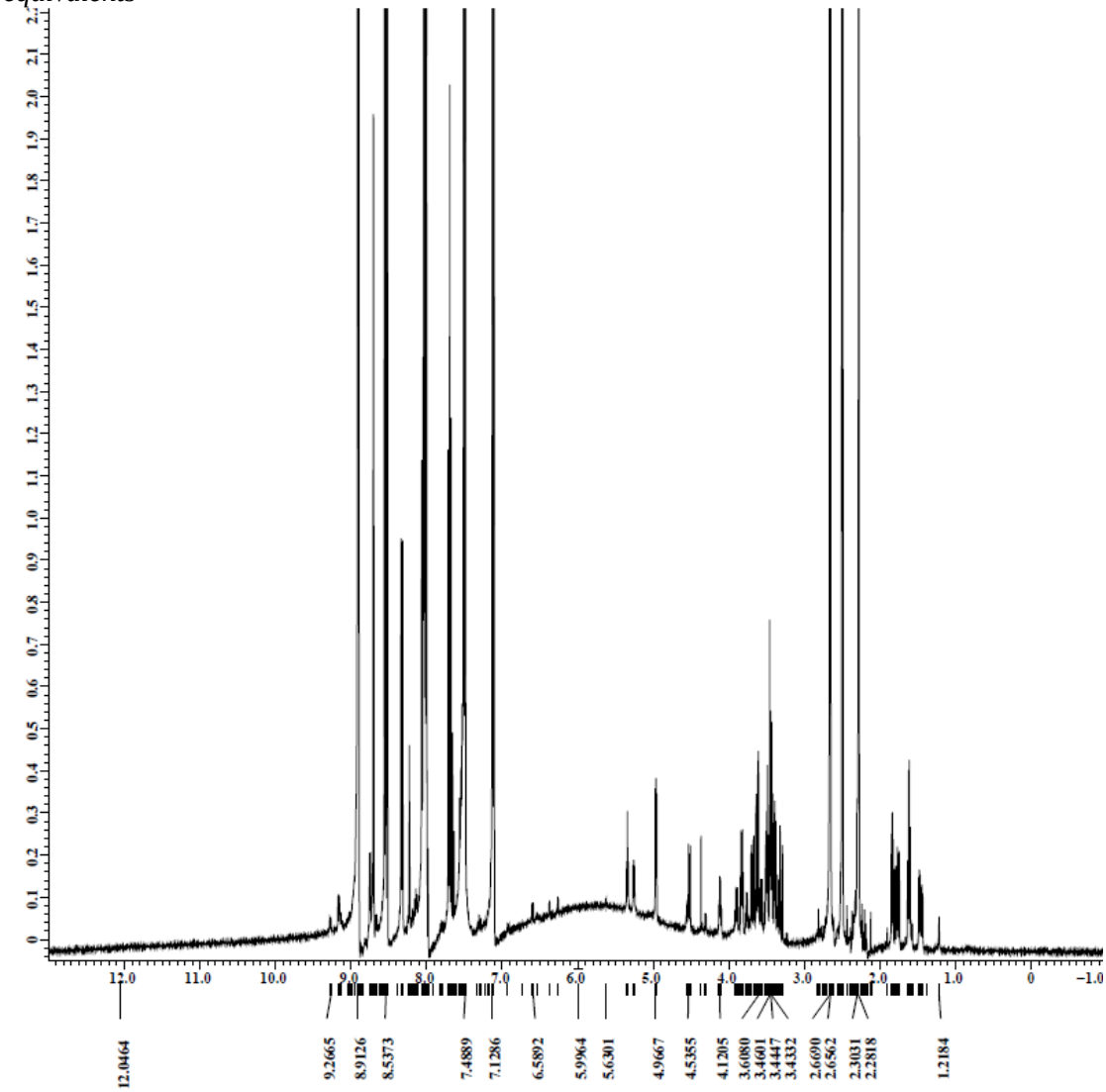


Figure A2.26: *^{1}H Nuclear Magnetic Resonance spectra for tosylation reaction crude 1.2 equivalents*

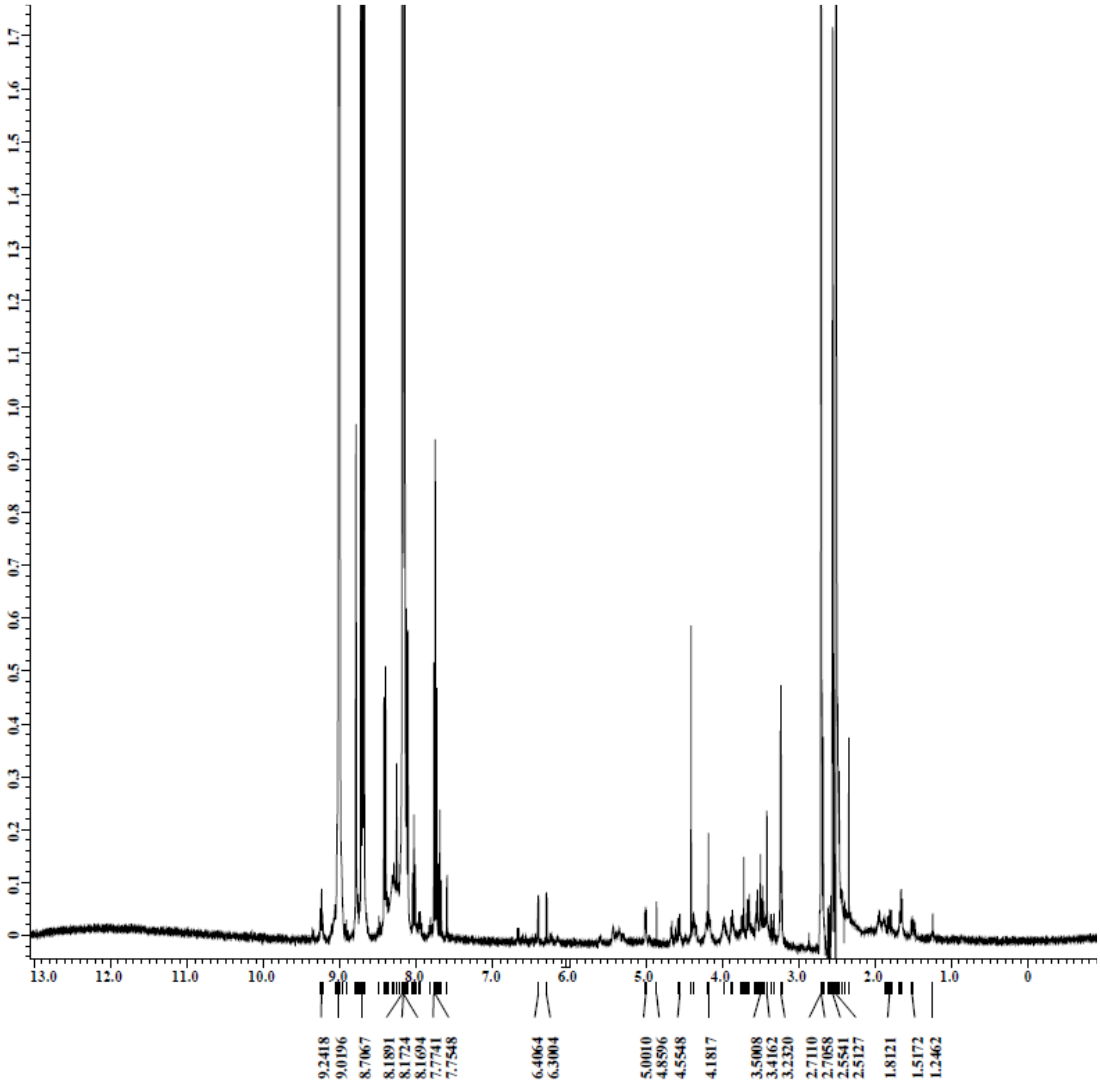


Figure A2.27: *^{1}H Nuclear Magnetic Resonance spectra for mesylation reaction 4 crude 1.2 equivalents*

2-hydroxy-dG

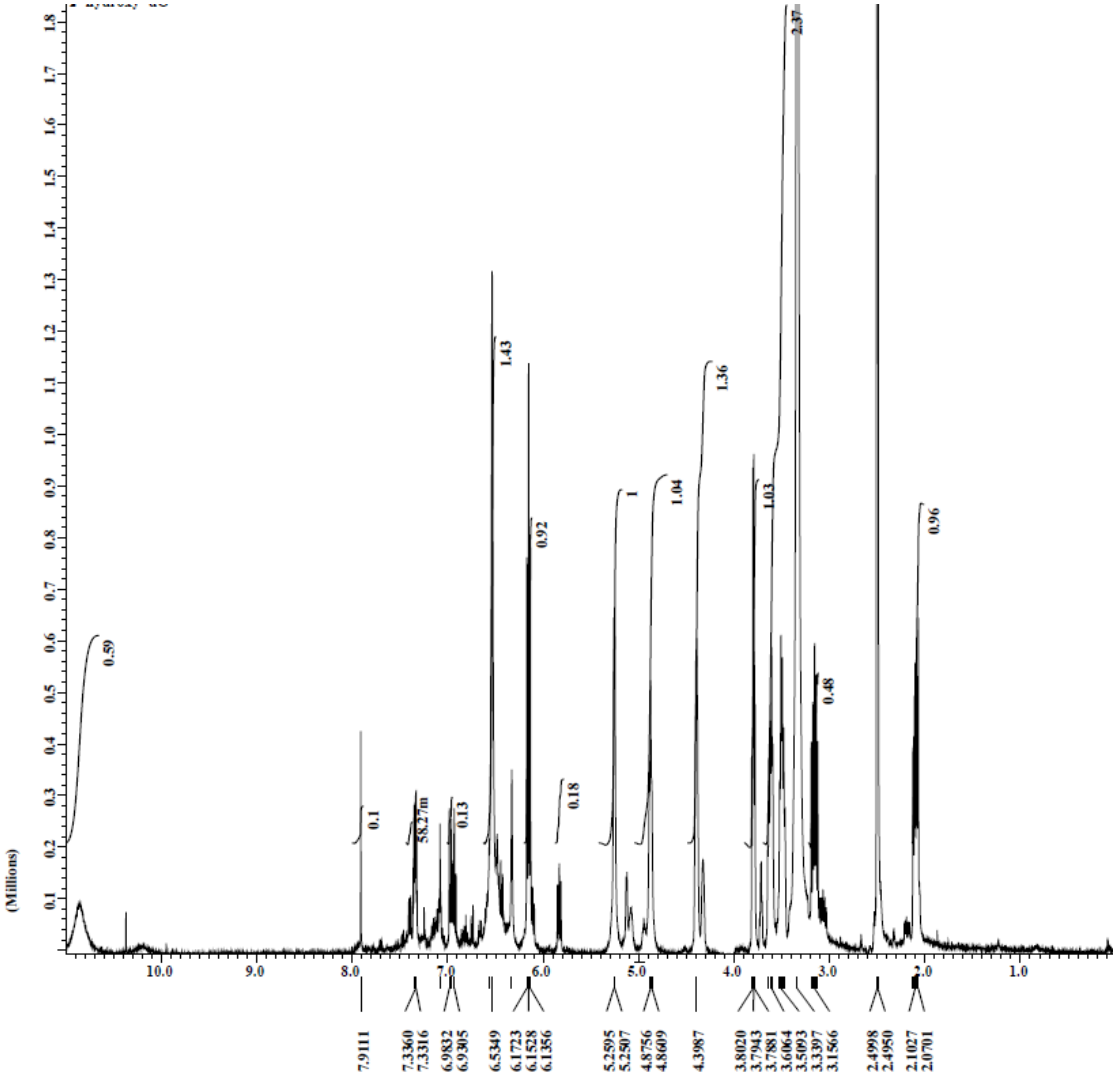


Figure A2.28: *^{1}H Nuclear Magnetic Resonance spectra of 2-hydroxyphenyl-2`-deoxyguanosine crude*

Appendix 3

Section 1: X-ray of Crystal structure of 4-pyrazole-dG

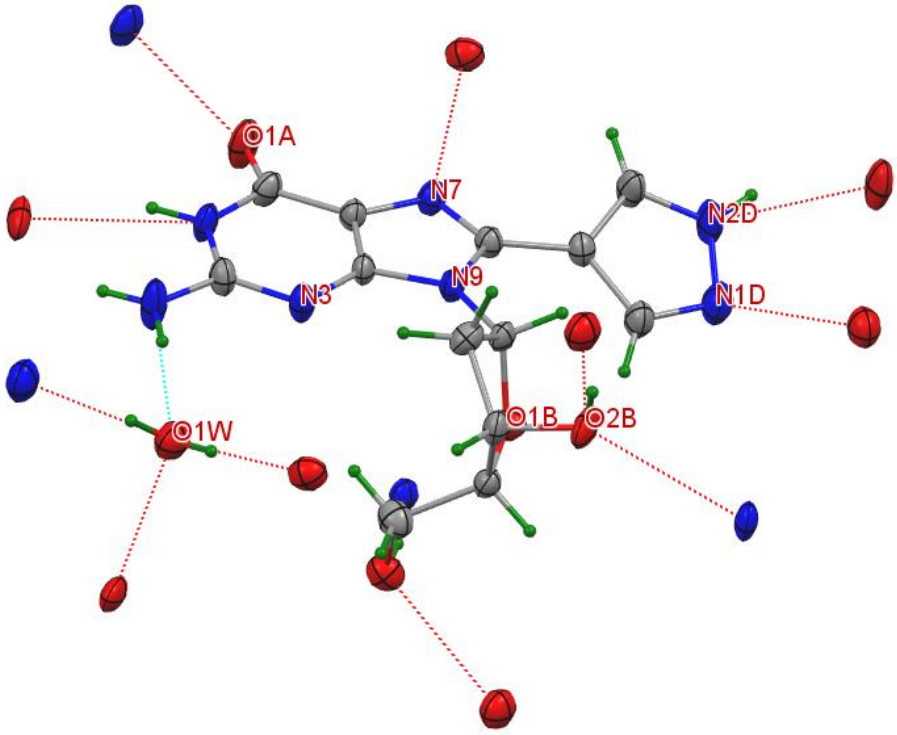


Figure A3.1: *X-ray crystal structure of 4-Pyrazole-dG*

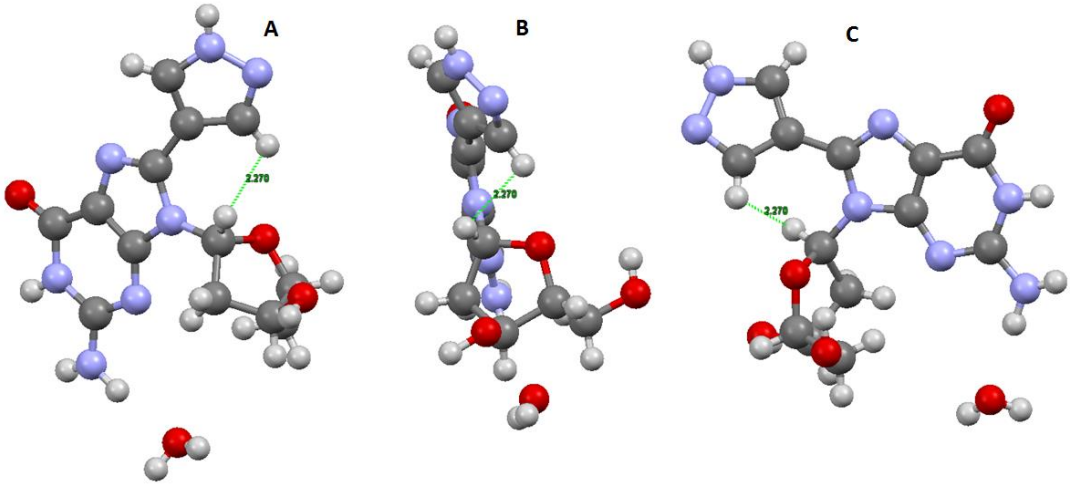


Figure A3.2: *X-ray crystal structure of 4-Pyrazole-dG, rotated on its Z-axis, displaying the front, back and side-on view of the molecule. The X-ray image displays close proximity between the H1`and the 3H`` protons as outlined in 3.4.1.3.*

Appendix 4

Mass spectrometry

8-(3-Acetylphenyl)-2'-deoxyguanosine, (3-Ac-dG)

Full range spectrum Elemental Composition *m/z* 386: The experimental molecular formula is consistent with the proposed molecular formula within ±5PPM mass error (typically required for publication or patent).

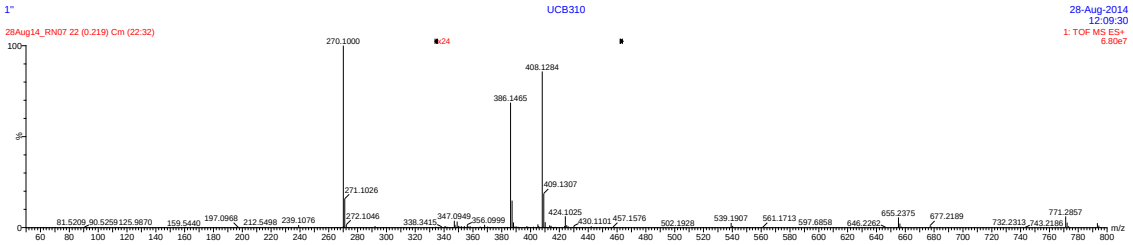


Figure A4.1: *Mass spectrum for 8-(3-Acetylphenyl)-2'-deoxyguanosine, (3-Ac-dG)*

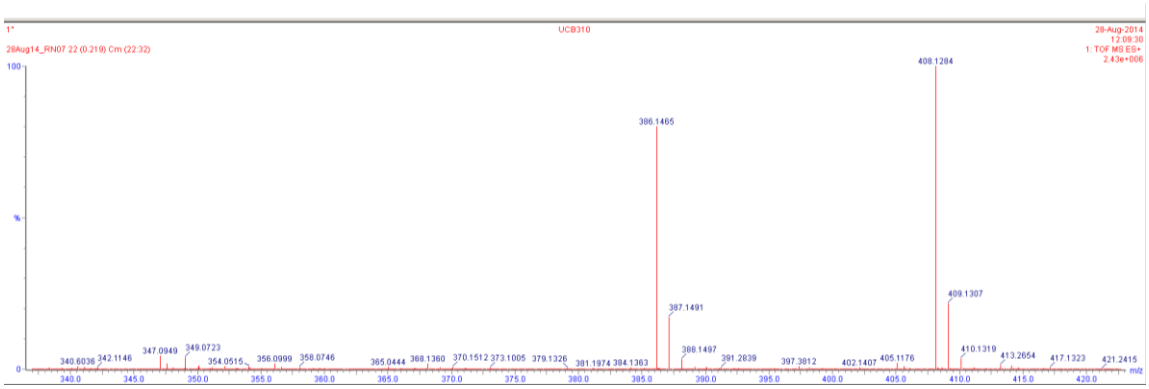


Figure A4.2: *Mass spectrum for 8-(3-Acetylphenyl)-2'-deoxyguanosine, (3-Ac-dG)*

3-hydroxy-dG

Full range spectrum Elemental Composition *m/z* 360 The experimental molecular formula is consistent with the proposed molecular formula within ±5PPM mass error (typically required for publication or patent).

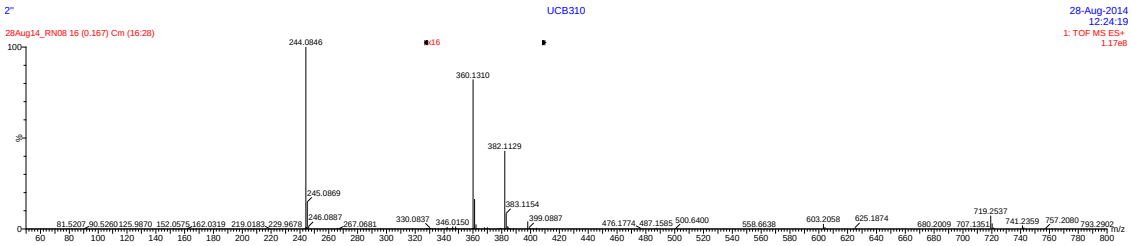


Figure A4.3: *Mass spectrum for 3-hydroxy-dG*

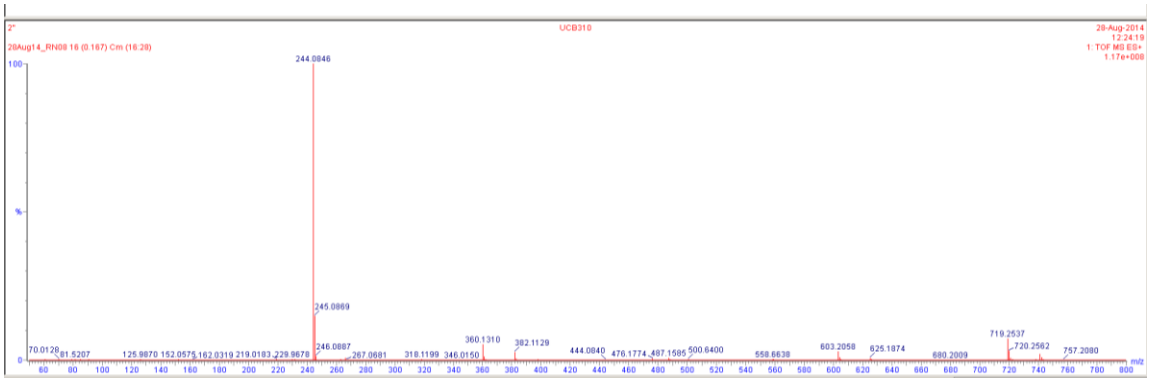


Figure A4.4: *Mass spectrum for 3-hydroxy-dG*

<u>*4-hydroxy-dG*</u>

Full range spectrum Elemental Composition *m/z* 360 The experimental molecular formula is consistent with the proposed molecular formula within ±5PPM mass error (typically required for publication or patent).

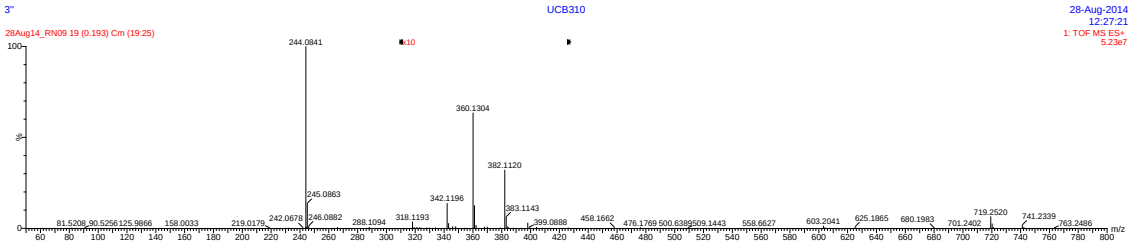


Figure A4.5: *Mass spectrum for 4-hydroxy-dG*

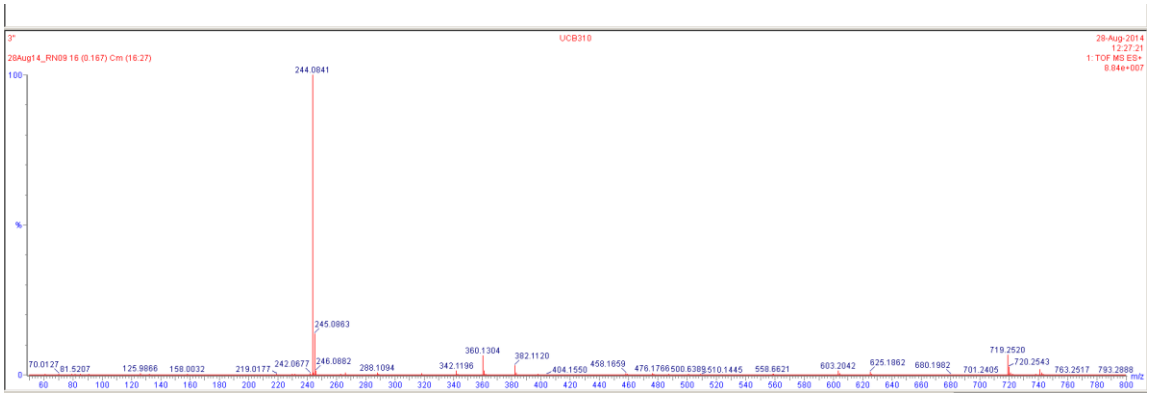


Figure A4.6: *Mass spectrum for 4-hydroxy-dG*

4-pyrazole-dG

Full range spectrum Elemental Composition *m/z* 334 The experimental molecular formula is consistent with the proposed molecular formula within ±5PPM mass error (typically required for publication or patent).

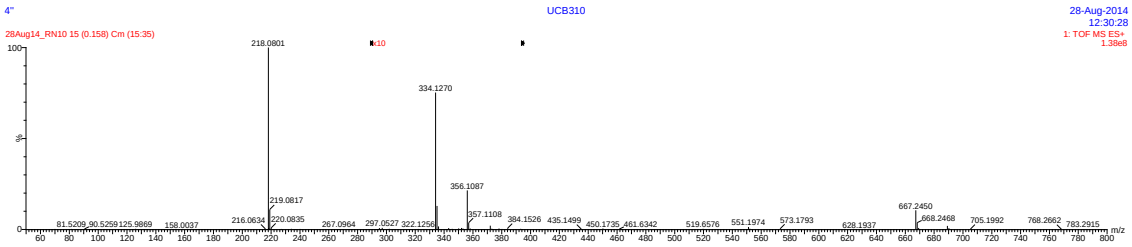


Figure A4.7: *Mass spectrum for 4-pyrazole-dG*

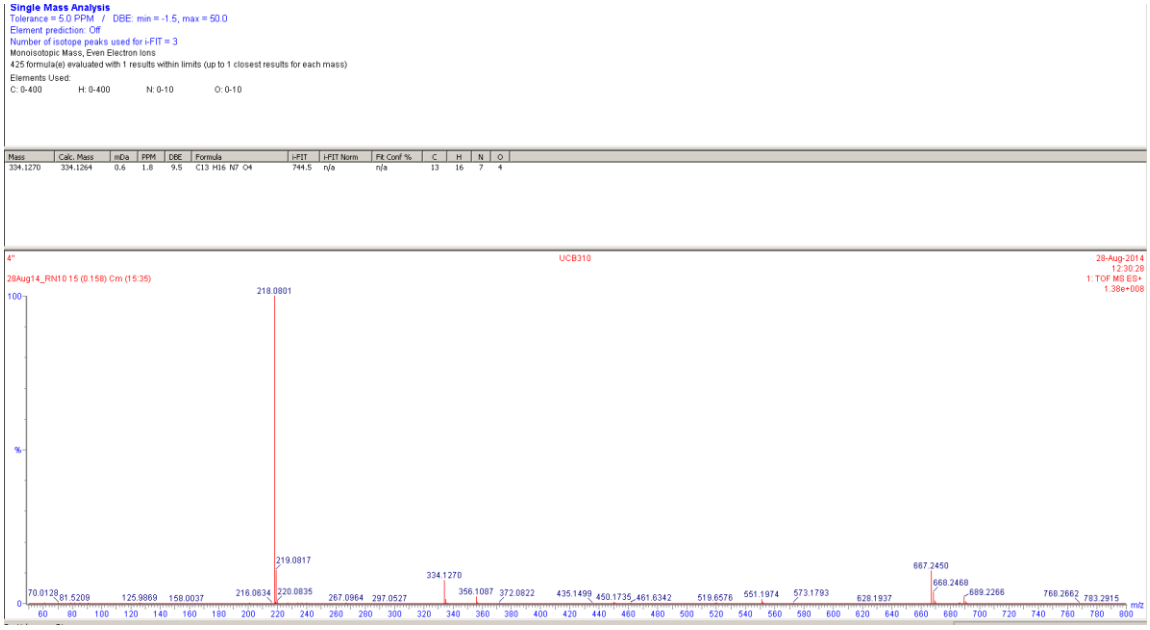


Figure A4.8: *Mass spectrum for 4-pyrazole-dG*

Full range spectrum Elemental composition *m/z* 334 The experimental molecular formula is consistent with the proposed molecular formula within ±5PPM mass error (typically required for publication or patent).

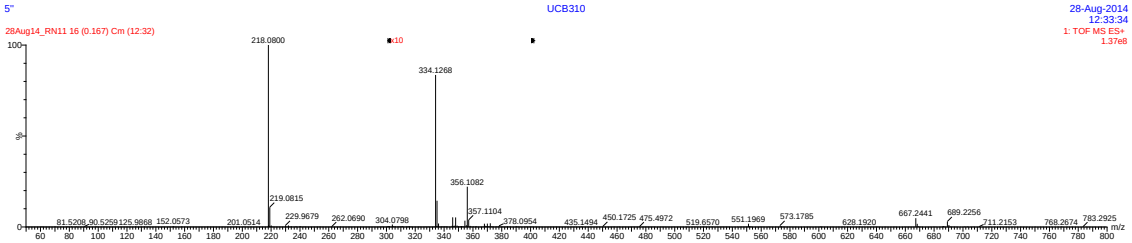


Figure A4.9: *Mass spectrum for 5-pyrazole-dG*

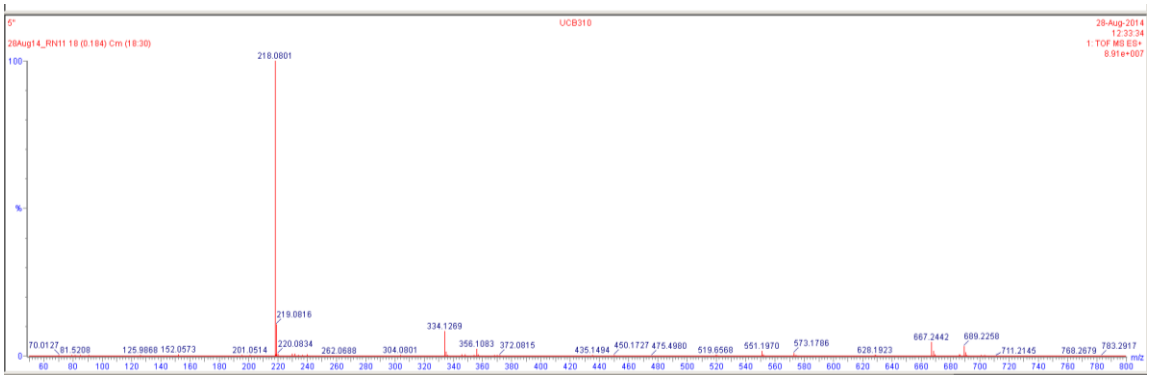


Figure A4.10: *Mass spectrum for 5-pyrazole-dG*

Appendix 5

Section 1: Explanation for method of calculation of Dual Polarisation Interferometry display data

The DPI display data is presented in Figures 2.1-2.4 as radians. This unit is calculated by the DPI Analight software version 1.7.The method of calculation is explained below in an extract from the paper **entitled '**A new quantitative optical biosensor for protein characterisation'

"The phase changes of interest, $\Delta\varphi$, involve alterations in the effective refractive index, *Ns*, of the mode in the upper, (sensing) waveguide. The effective refractive index of the lower (reference) mode, *Nr*, is unaffected by changes occurring at the surface since the evanescent field of this mode decays rapidly in the region between the two guiding layers. Direct measurement of $\Delta\varphi$ is possible by continuously monitoring the relative phase position of the fringe pattern by performing a Fourier transformation relating intensity to position. Thus, we can easily convert the experimental data to changes in effective refractive index since the path length is fixed. The lower mode's effective refractive index remains unchanged, therefore, the measured change can be related solely to that experienced by the upper waveguide mode.

Having obtained $\Delta\varphi$, it is then possible to calculate a range of layer thicknesses and refractive index values. In order to determine a final solution rapidly we employ a successive approximation method to search for the optimum refractive index value and thickness of the thin layers growing at the surface of the waveguide."

"The absolute effective refractive index of a waveguide mode is found by solving the equations of electromagnetism for a system of uniform multiple dielectric layers in which the fields in the semi-infinite bounding layers have exponentially decaying solutions. It is assumed for the purposes of calculation that the layers are isotropic in nature. The parameters required are the refractive index and thickness of each layer (except the bounding layers where index only is specified) for each of the two allowed states of polarisation (TE, and TM). Provided the input information is complete, an effective refractive index value for each mode is obtained which is representative of the distribution of optical power amongst the layers. This is the starting point for any analysis of the measured phase changes.

If a new layer is introduced to (or removed from) the system it will alter the effective refractive index of the mode. For each of the two polarisation states, the new effective refractive index can satisfy a continuous range of thickness and refractive index values for the new layer. However, there will only be one unique combination that satisfies the effective refractive index of both the TE and TM modes simultaneously."

To summarise the DPI instrument uses Fourier transform algorithms in conjunction with refractive Index measurements of sample materials, adhered to the surface of an Anachip to produce units in radians. The radian units are graphed against time in seconds.

Section 2: Dual Polarisation Interferometry display data for DPI experiments # 1-4

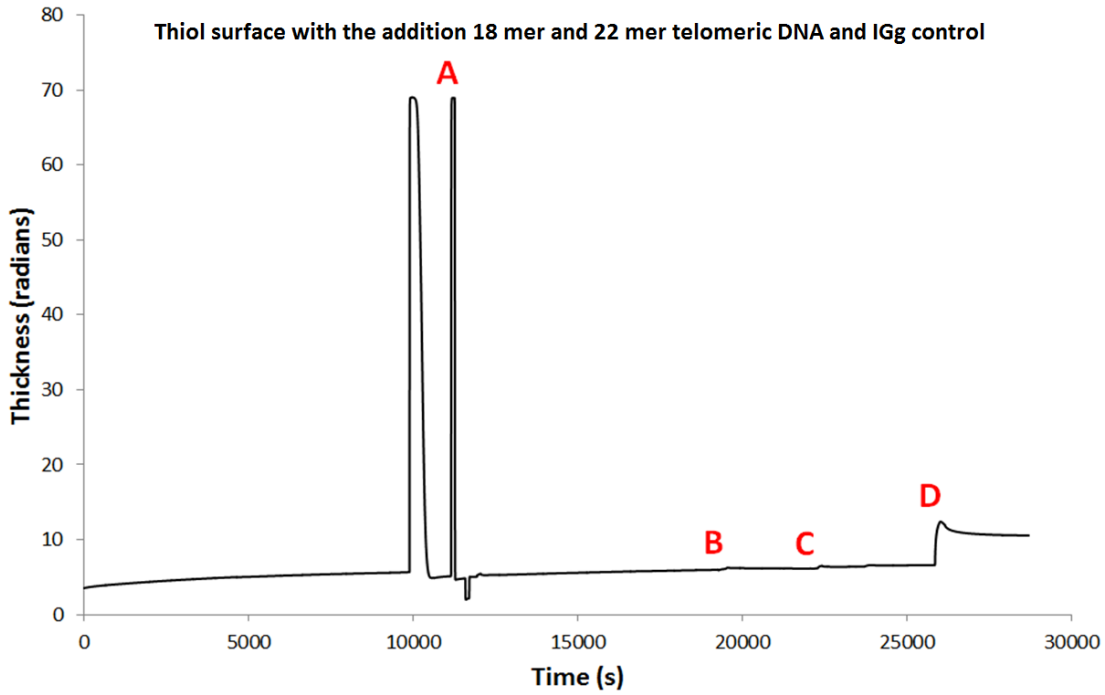


Figure A5.1: (DPI experiment # 1) *A thiol-coated Anachip surface was calibrated using high purity ethanol and water as standards (A).100 µM 18-mer, thiolated, cy5-labelled DNA in H_2O was injected onto the Anachip surface and allowed to incubate for 2 hours (B). 100 µM 22-mer thiolatedcy5-labelled DNA in H_2O was injected onto the Anachip surface and allowed to incubate for 2 hours (C). A solution of 10 µM IgG protein in H_2O was injected onto the Anachip surface (D).*

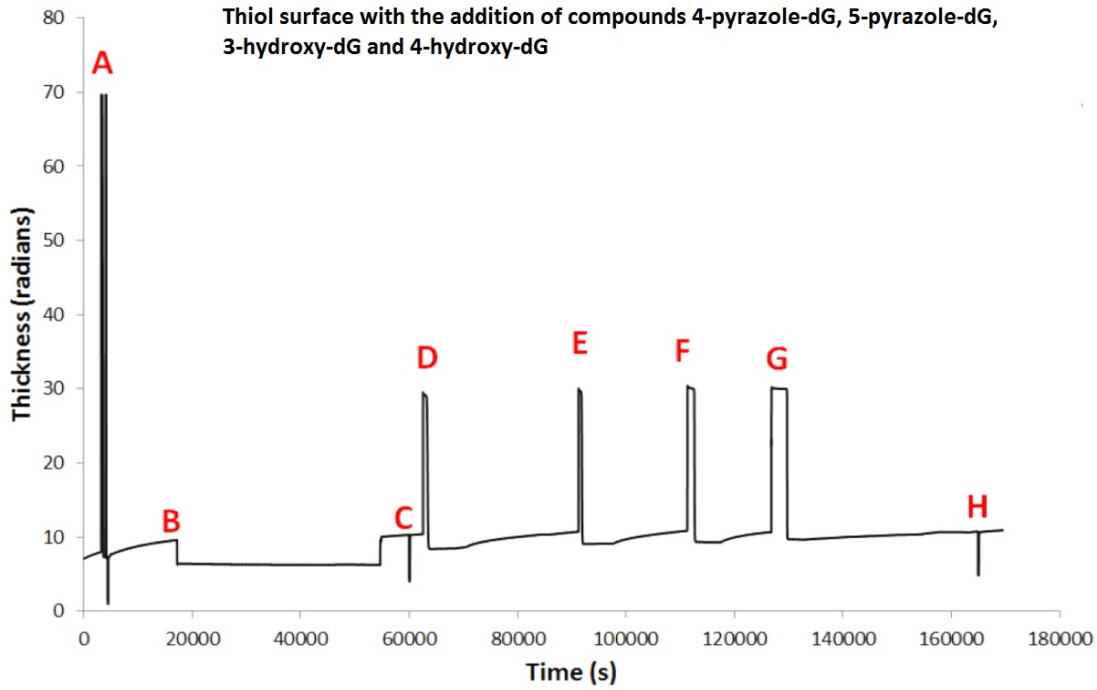


Figure A5.2: (DPI experiment # 2) *A thiol-coated Anachip surface was calibrated using high purity ethanol and water as standards (A). 1µM 30-mer, thiolated, cy5-labelled DNA in H_2O was injected onto the Anachip surface and allowed to incubate for 10 hours (B). 5 mM potassium iodide in H_2O was injected onto the Anachip surface (C). Four solutions of acetonitrile and water in a 1:1 mixture, containing 100 µM potassium iodide and 100 µM of the compounds 3-hydroxy-dG (D), 4-pyrazole-dG (E), 5-pyrazole-dG (F) and 4-hydroxy-dG (G) were injected onto the surface of the Anachip. A solution of 5mM sucrose in H_2O (H) was injected onto the Anachip surface.*

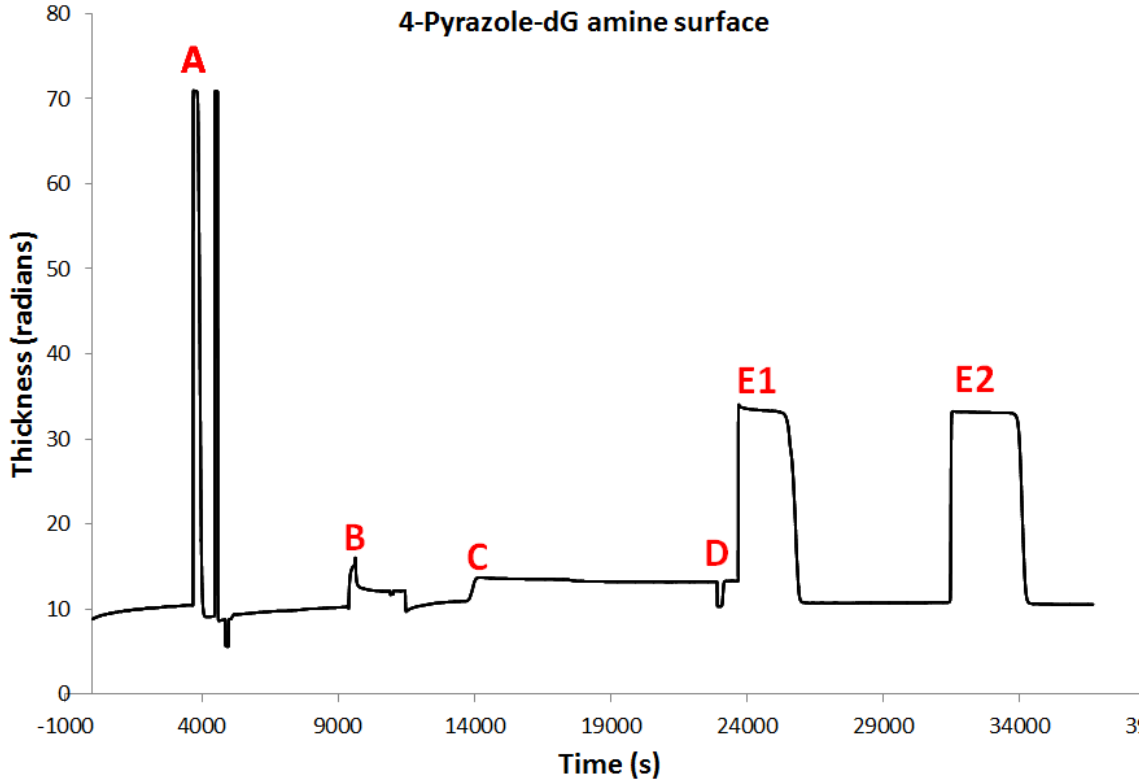


Figure A5.3: (DPI experiment # 3) *Amine-coated Anachip surface was calibrated using high purity ethanol and water as standards (A). A solution of 5 mM Sulfo-SMCC in H_2O was injected onto the Anachip surface (B). 1µM 30-mer, thiolated, cy5-labelled DNA in H_2O was injected onto the Anachip surface 5 mM potassium iodide in H_2O was injected onto the Anachip surface (C). An aqueous 5 mM solution of potassium iodide was injected onto the Anachip surface (D). A solution of acetonitrile and water in a 1:1 mixture, containing 100 µM 4-pyrazole-dG and 100 µM of the potassium iodide was injected onto the surface of the Anachip (E), the 100 µM 4-pyrazole-dG solution was injected again (E2).*

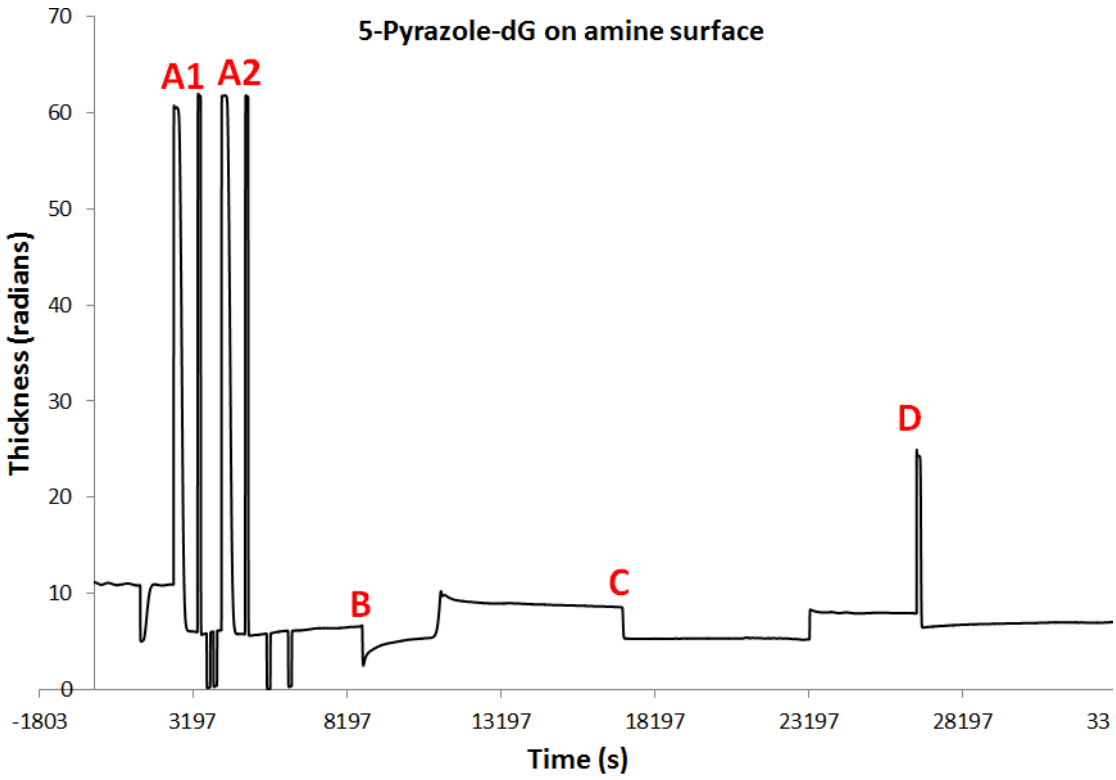


Figure A5.4: (DPI experiment # 4) *An amine-coated Anachip surface was repeatedly calibrated using high purity ethanol and water as standards, displayed as peaks (A1) and (A2). A solution of 5 mM Sulfo-SMCC in H_2O was injected onto the Anachip surface (B). 1 µM 30-mer thiolated-cy5-labelled DNA in H_2O was injected onto the Anachip surface (C). A solution of acetonitrile and water in a 1:1 mixture, containing 100 µM 5-pyrazole-dG and 100 µM of the potassium iodide was injected onto the surface of the Anachip (D).*

References

1. O'Sullivan, R.J. and, Karlseder, J. Telomeres: Protecting chromosomes against genome instability. *Nature Reviews Molecular Cell Biology*. 2010;11(3):171-181. doi: 10.1038/nrm2848.

2. Blackburn E. and, Gall, J. Tandemly repeated sequence at termini of extrachromosomal ribosomal-rna genes in tetrahymena. *Journal of Molecular Biology*. 1978;120(1):33-53. doi: 10.1016/0022-2836(78)90294-2.

3. Makarov, V., Hirose, Y., Langmore, J. and, Long, G. Tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*. 1997;88(5):657-666. doi: 10.1016/S0092-8674(00)81908-X.

4. Blasco, M. Telomeres and human disease: Ageing, cancer and beyond. *Nature Reviews Genetics*. 2005;6(8):611-622. doi: 10.1038/nrg1656.

5. Nandakurnar, J. and, Cech, T.R. Finding the end: Recruitment of telomerase to telomeres. *Nature Reviews Molecular Cell Biology*. 2013;14(2):69-82. doi: 10.1038/nrm3505.

6. Hayflick, L. Limited *in vitro* lifetime of human diploid cell strains. *Experimental Cell Research*. 1965;37(3):614-636. doi: 10.1016/0014-4827(65)90211-9.

7. Greider, C. and, Blackburn, E. The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with 2 kinds of primer specificity. *Cell*. 1987;51(6):887-898. doi: 10.1016/0092-8674(87):90576-9.

8. Nakamura, T., Morin, G., Chapman, K., et al. Telomerase catalytic subunit homologs from fission yeast and human. *Science*. 1997;277(5328):955-959. doi: 10.1126/science.277.5328.955.

9. Venteicher, A.S., Abreu, E.B., Meng, Z., et al. A human telomerase holoenzyme protein required for cajal body localization and telomere synthesis. *Science*. 2009;323(5914):644-648. doi: 10.1126/science.1165357.

10. Venteicher, A.S., Meng, Z., Mason, P.J., Veenstra, T.D. and, Artandi, S.E. Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell*. 2008;132(6):945-957. doi: 10.1016/j.cell.2008.01.019.

11. De Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development*. 2005;19(18):2100-2110. doi: 10.1101/gad.1346005.

12. Kirwan, M. and, Dokal, I. Dyskeratosis congenita, stem cells and telomeres. *Biochimica et Biophysica Acta - Molecular Basis of Disease*. 2009;1792(4):371-379. doi: 10.1016/j.bbadis.2009.01.010.

13. Kim, N., Piatyszek, M., Prowse K., et al. Specific association of human telomerase activity with immortal cells and cancer. *Science*. 1994;266(5193):2011-2015. doi: 10.1126/science.7605428.

14. Heidinger, B.J., Blount, J.D., Boner, W., Griffiths, K., Metcalfe, N.B. and, Monaghan, P. Telomere length in early life predicts lifespan. *Proceedings of the National Academy of Sciences U S A*. 2012;109(5):1743-1748. doi: 10.1073/pnas.1113306109.

15. Bodnar, A., Ouellette, M., Frolkis, M., et al. Extension of life-span by introduction of telomerase into normal human cells. *Science*. 1998;279(5349):349-352. doi: 10.1126/science.279.5349.349.

16. Jaskelioff, M., Muller, F.L., Paik, J., et al. Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice. *Nature*. 2011;469(7328):102-U1700. doi: 10.1038/nature09603.

17. Joeng, K., Song, E., Lee, K. and, Lee, J. Long lifespan in worms with long telomeric DNA. *Nature Genetics*. 2004;36(6):607-611. doi: 10.1038/ng1356.

18. Henderson, E., Hardin, C., Walk, S., Tinoco, I. and, Blackburn, E. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine guanine base-pairs. *Cell*. 1987;51(6):899-908. doi: 10.1016/0092-8674(87)90577-0.

19. Sun, D., Thompson, B., Cathers, B., et al. Inhibition of human telomerase by a G-quadruplex-interactive compound. *Journal of Medicinal Chemistry*. 1997;40(14):2113-2116. doi: 10.1021/jm970199z.

20. Bugaut, A., Rodriguez, R., Kumari, S., Hsu, S.D. and, Balasubramanian, S. Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex. *Organic and Biomolecular Chemistry*. 2010;8(12):2771-2776. doi: 10.1039/c002418j.

21. Burger, A., Dai, F., Schultes, C., et al. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Research*. 2005;65(4):1489-1496. doi: 10.1158/0008-5472.CAN-04-2910.

22. Pilch, D.S., Barbieri, C.M., Rzuczek, S.G., LaVoie, E.J. and, Rice, J.E. Targeting human telomeric G-quadruplex DNA with oxazole-containing macrocyclic compounds. *Biochimie*. 2008;90(8):1233-1249. doi: 10.1016/j.biochi.2008.03.011.

23. Salvati, E., Leonetti, C., Rizzo, A., et al. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *Journal of Clinical Investigation*. 2007;117(11):3236-3247. doi: 10.1172/JCI32461.

24. Rawal, P., Kummarasetti, V., Ravindran, J., et al. Genome-wide prediction of G4 DNA as regulatory motifs: Role in *Escherichia coli* global regulation. *Genome Research*. 2006;16(5):644-655. doi: 10.1101/gr.4508806.

25. Mullen, M.A., Olson, K.J., Dallaire, P., Major, F., Assmann, S.M. and, Bevilacqua, P.C. RNA G-quadruplexes in the model plant species *Arabidopsis thaliana*: Prevalence and possible functional roles. *Nucleic Acids Research*. 2010;38(22):8149-8163. doi: 10.1093/nar/gkq804.

26. Yadav, V.K., Abraham, J.K., Mani, P., Kulshrestha, R. and, Chowdhury, S. QuadBase: Genome-wide database of G4 DNA - occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. *Nucleic Acids Research*. 2008;36:D381-D385. doi: 10.1093/nar/gkm781.

27. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. and, Lipps, H. Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*. *Nature Structural and Molecular Biology*. 2005;12(10):847-854. doi: 10.1038/nsmb982.

28. Giraldo, R., Suzuki, M., Chapman, L. and, Rhodes, D. Promotion of parallel DNA quadruplexes by a yeast telomere binding-protein - a circular-dichroism study. *Proceedings of the National Academy of Sciences U S A*. 1994;91(16):7658-7662. doi: 10.1073/pnas.91.16.7658.

29. Mergny, J., Phan, A. and, Lacroix, L. Following G-quartet formation by UV-spectroscopy.*FEBS Letters*. 1998;435(1):74-78. doi: 10.1016/S0014-5793(98)01043-6.

30. Vairamani, M. and, Gross, M. G-quadruplex formation of thrombin-binding aptamer detected by electrospray ionization mass spectrometry. *Journal of the American Chemical Society*. 2003;125(1):42-43. doi: 10.1021/ja0284299.

31. Phillips, K., Dauter, Z., Murchie, A., Lilley, D. and, Luisi, B.The crystal structure of a parallel-stranded guanine tetraplex at 0.95 angstrom resolution. *Journal of Molecular Biology*. 1997;273(1):171-182. doi: 10.1006/jmbi.1997.1292.

32. Patel, P., Ainavarapu, R. and, Hosur, R. NMR studies on truncated sequences of human telomeric DNA: Observation of a novel A-tetrad. *Nucleic Acids Research*. 1999;27(19):3836-3843. doi: 10.1093/nar/27.19.3836.

33. Lipps, H.J. and, Rhodes, D. G-quadruplex structures: *In vivo* evidence and function. *Trends in Cell Biology*. 2009;19(8):414-422. doi: 10.1016/j.tcb.2009.05.002.

34. Hardin, C., Henderson, E., Watson, T. and, Prosser, J. Monovalent cation induced structural transitions in telomeric DNAs - G-DNA folding intermediates. *Biochemistry (N Y)*. 1991;30(18):4460-4472. doi: 10.1021/bi00232a013.

35. Yafe, A., Etzioni, S., Weisman-Shomer, P. and, Fry, M. Formation and properties of hairpin and tetraplex structures of guanine-rich regulatory sequences of muscle-specific genes. *Nucleic Acids Research*. 2005;33(9):2887-2900. doi: 10.1093/nar/gki606.

36. Patel, D.J., Phan, A.T. and, Kuryavyi, V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Research*. 2007;35(22):7429-7455. doi: 10.1093/nar/gkm711.

37. Fedoroff, O., Salazar, M., Han, H., Chemeris, V., Kerwin, S. and, Hurley, L. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry (N Y)*. 1998;37(36):12367-12374. doi: 10.1021/bi981330n.

38. Read, M. and, Neidle, S. Structural characterization of a guanine-quadruplex ligand complex. *Biochemistry (N Y)*. 2000;39(44):13422-13432. doi: 10.1021/bi001584k.

39. Clark, G.R., Pytel, P.D. and, Squire, C.J. The high-resolution crystal structure of a parallel intermolecular DNA G-4 quadruplex/drug complex employing syn glycosyl linkages. *Nucleic Acids Research*. 2012;40(12):5731-5738. doi: 10.1093/nar/gks193.

40. Clark, G., Pytel, P., Squire, C. and, Neidle, S. Structure of the first parallel DNA quadruplex-drug complex. *Journal of the American Chemical Society*. 2003;125(14):4066-4067. doi: 10.1021/ja0297988.

41. Nicoludis, J.M., Miller, S.T., Jeffrey, P.D., et al. Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *Journal of the American Chemical Society*. 2012;134(50):20446-20456. doi: 10.1021/ja3088746.

42. Green, J.J., Ladame, S., Ying, L., Klenerman, D. and, Balasubramanian, S. Investigating a quadruplex-ligand interaction by unfolding kinetics. *Journal of the American Chemical Society*. 2006;128(30):9809-9812. doi: 10.1021/ja0615425.

43. Tendian, S. and, Parker, W. Interaction of deoxyguanosine nucleotide analogs with human telomerase. *Molecular Pharmacology*. 2000;57(4):695-699.

44. Fletcher, T., Cathers, B., Ravikumar, K., Mamiya, B. and, Kerwin, S. Inhibition of human telomerase by 7-deaza-2 '-deoxyguanosine nucleoside triphosphate analogs: Potent inhibition by 6-thio-7-deaza-2 '-deoxyguanosine 5 '-triphosphate. *Bioorganic Chemistry*. 2001;29(1):36-55. doi: 10.1006/bioo.2000.1194.

45. Maidak, B., Olsen, G., Larsen, N., Overbeek, R., McCaughey, M. and, Woese, C. The ribosomal database project (RDP). *Nucleic Acids Research*. 1996;24(1):82-85. doi: 10.1093/nar/24.1.82.

46. Kostadinov, R., Malhotra, N., Viotti, M., Shine, R., D'Antonio, L. and, Bagga, P. GRSDB: A database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences. *Nucleic Acids Research*. 2006;34:D119-D124. doi: 10.1093/nar/gkj073.

47. Kikin, O., Zappala, Z., D'Antonio, L. and, Bagga, P.S. GRSDB2 and GRS_UTRdb: Databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs. *Nucleic Acids Research*. 2008;36:D141-D148. doi: 10.1093/nar/gkm982.

48. Kaucher, M., Harrell, W. and, Davis, J. A unimolecular G-quadruplex that functions as a synthetic transmembrane Na^+ transporter. *Journal of the American Chemical Society*. 2006;128(1):38-39. doi: 10.1021/ja056888e.

49. Cheng, X., Liu, X., Bing, T., Cao, Z. and, Shangguan, D. General peroxidase activity of G-quadruplex-hemin complexes and its application in ligand screening. *Biochemistry (N Y)*. 2009;48(33):7817-7823. doi: 10.1021/bi9006786.

50. Zhao, Y., Kan, Z., Zeng, Z., Hao, Y., Chen, H. and, Tan, Z. Determining the folding and unfolding rate constants of nucleic acids by biosensor. Application to telomere G-quadruplex. *Journal of the American Chemical Society* 2004;126(41):13255-13264. doi: 10.1021/ja048398c.

51. Li, T., Dong, S. and, Wang, E. Label-free colorimetric detection of aqueous mercury ion (Hg^{2+}) using Hg^{2+}-modulated G-quadruplex-based DNAzymes. *Analytical Chemistry*. 2009;81(6):2144-2149. doi: 10.1021/ac900188y.

52. Yao, Y., Wang, Q., Hao, Y. and, Tan, Z. An exonuclease I hydrolysis assay for evaluating G-quadruplex stabilization by small molecules.*Nucleic Acids Research*. 2007;35(9):e68. doi: 10.1093/nar/gkm194.

53. Lam, E.Y.N, Beraldi, D., Tannahill, D. and, Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nature Communications*. 2013;4:1796. doi: 10.1038/ncomms2792.

54. Gubala, V., Betancourt, J. and, Rivera, J. Expanding the hoogsteen edge of 2 '-deoxyguanosine: Consequences for G-quadruplex formation. *Organic Letters*. 2004;6(25):4735-4738. doi: 10.1021/ol048013v.

55. Rivera-Sanchez, Md.C., Andujar-de-Sanctis, I., Garcia-Arriaga, M., Gubala, V., Hobley, G. and, Rivera, J.M. Walking a supramolecular tightrope: A self-assembled dodecamer from an 8-aryl-2 '-deoxyguanosine derivative. *Journal of the American Chemical Society*. 2009;131(30):10403-+. doi: 10.1021/ja9040384.

56. Wang, Y., Delossantos, C., Gao, X., Greene, K., Live, D. and, Patel, D. Multinuclear nuclear-magnetic-resonance studies of Na cation-stabilized complex formed by D(g-G-T-T-T-T-C-G-G) in solution - implications for G-tetrad structures. *Journal of Molecular Biology*. 1991;222(3):819-832. doi: 10.1016/0022-2836(91)90513-6.

57. Kolb, H., Finn, M. and, Sharpless, K. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie*. 2001;40(11):2004-+. doi: 10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

58. Tornoe, C., Christensen, C. and, Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *Journal of Organic Chemistry*. 2002;67(9):3057-3064. doi: 10.1021/jo011148j.

59. Nakata, T., Nomura, S. and, Matsukura, H. Stereoselective synthesis of six- and seven-membered ether rings based on the ring expansion. *Tetrahedron Letters*. 1996;37(2):213-216. doi: 10.1016/0040-4039(95)02131-0.

60. Collier, A. and, Wagner, G.K., Suzuki-Miyaura cross-coupling of unprotected halopurine nucleosides in water-influence of catalyst and cosolvent. *Synthetic Communications.*2006;36(24):3713-3721. doi: 10.1080/10916460600946139.

61. Miyaura, N., Yamada, K. and, Suzuki, A. New stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides.*Tetrahedron Lett*. 1979;20(36):3437-3440. doi: 10.1016/S0040-4039(01)95429-2.

62. Miyaura, N. and, Suzuki, A. Stereoselective synthesis of arylated (E)-alkenes by the reaction of alk-1-enylboranes with aryl halides in the presence of palladium catalyst. *Chemical Communications.* 1979(19):866-867. doi: 10.1039/c39790000866.

63. Stille, J. and, Lau, K. Mechanisms of oxidative addition of organic halides to group-8 transition-metal complexes. *Accounts of Chemical Research*. 1977;10(12):434-442. doi: 10.1021/ar50120a002.

64. Casado, A. and, Espinet, P. On the configuration resulting from oxidative addition of RX to pd(PPh3)(4) and the mechanism of the cis-to-trans isomerization of [PdRX(PPh3)(2)] complexes (R equals aryl, X equals halide). *Organometallics*. 1998;17(5):954-959. doi: 10.1021/om9709502.

65. Inanaga, J., Hiratak, K., Saeki, H., Katsuki, T. and, Yamaguchi, M. Rapid esterification by means of mixed anhydride and its application to large-ring lactonization. *Bulletin of the Chemical Society of Japan*. 1979;52(7):1989-1993. doi: 10.1246/bcsj.52.1989.

66. Kolb, H. and, Sharpless, K. The growing impact of click chemistry on drug discovery.*Drug Discovery Today*. 2003;8(24):1128-1137. doi: 10.1016/S1359-6446(03)02933-7.

67. Appel, R. Tertiary phosphane-tetrachloromethane, a versatile reagent for chlorination, dehydration, and P-N linkage. *Angewandte Chemie International Edition* . 1975;14(12):801-811. doi: 10.1002/anie.197508011.

68. Davisson, V., Woodside, A., Neal, T., Stremler, K., Muehlbacher, M. and, Poulter, C. Phosphorylation of isoprenoid alcohols. *Journal of Organic Chemistry*. 1986;51(25):4768-4779. doi: 10.1021/jo00375a005.

69. Tamaddon, F., Nasiri, A. And, Farokhi, S. CsF-celite as an efficient heterogeneous catalyst for sulfonylation and desulfonylation of heteroatoms. *Catalysis Communications*. 2011;12(15):1477-1482. doi: 10.1016/j.catcom.2011.04.005.

70. Frye, E.C., O'Connor, C.J., Twigg, D.G., et al. Palladium-catalysed cross-coupling of vinyldisiloxanes with benzylic and allylic halides and sulfonates. *Chemistry - A European Journal*. 2012;18(28):8774-8779. doi: 10.1002/chem.201200431.

74. Hudson, J.S., Brooks, S.C. and, Graves, D.E. Interactions of actinomycin D with human telomeric G-quadruplex DNA. *Biochemistry (N Y)*. 2009;48(21):4440-4447. doi: 10.1021/bi900203z.

75. Nagatoishi, S., Tanaka, Y. and, Tsumoto, K. Circular dichroism spectra demonstrate formation of the thrombin-binding DNA aptamer G-quadruplex under stabilizing-cation-deficient conditions (vol 352, pg 812, 2007). *Biochemical and Biophysical Research Communications*. 2007;354(3):837-838. doi: 10.1016/j.bbrc.2007.01.018.

76. Dai, J., Chen, D., Jones, R.A., Hurley, L.H. and, Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Research*. 2006;34(18):5133-5144. doi: 10.1093/nar/gkl610.

77. Phan, A. and, Mergny, J. Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. *Nucleic Acids Research*. 2002;30(21):4618-4625. doi: 10.1093/nar/gkf597.

78. Sacca, B., Lacroix, L. and, Mergny, J. The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Research*. 2005;33(4):1182-1192. doi: 10.1093/nar/gki257.

79. Costa, L., Kerkmann, M., Hartmann, G., et al. Structural studies of oligonucleotides containing G-quadruplex motifs using AFM. *Biochemical and Biophysical Research Communications.* 2004;313(4):1065-1072. doi: 10.1016/j.bbrc.2003.12.041.

80. Higashiyama, T. Novel functions and applications of trehalose. *Pure Applied Chemistry*. 2002;74(7):1263-1269. doi: 10.1351/pac200274071263.

81. Herne, T. and, Tarlov, M. Characterization of DNA probes immobilized on gold surfaces. *Journal of Organic Chemistry*. 1997;119(38):8916-8920. doi: 10.1021/ja9719586.

82. Boudjemline, A., Clarke, D.T., Freeman, N.J., Nicholson, J.M. and, Jones, G.R. Early stages of protein crystallization as revealed by emerging optical waveguide technology. *Journal of Applied Crystallography*. 2008;41:523-530. doi: 10.1107/S0021889808005098.

83. Le, N.C.H., Gubala, V., Gandhiraman, R.P., Daniels, S. And, Williams, D.E. Evaluation of different nonspecific binding blocking agents deposited inside poly(methyl methacrylate) microfluidic flow-cells. *Langmuir*. 2011;27(14):9043-9051. doi: 10.1021/la2011502.

84. Peterlinz, K., Georgiadis, R., Herne, T. and, Tarlov, M. Observation of hybridization and dehybridization of thiol-tethered DNA using two-color surface plasmon resonance spectroscopy. *Journal of American Chemistry Society*. 1997;119(14):3401-3402. doi: 10.1021/ja964326c.

85. Wingvist, A., Stromberg, R. Investigation on Condensing Agents for Phosphinate Ester Formation with Nucleoside 5′-Hydroxyl Functions. *European Journal of Organic Chemistry*. 2008;10;17505-1715. doi 10.1002/ejoc.200700528

86. Spain, E., Brennan, E., Keyes, T.E. and, Forster, R.J. Dual function metal nanoparticles: Electrocatalysis and DNA capture. *Electrochimica Acta*. 2014;128;61-66. doi 10.1016/j.electacta.2013.10.025

87. Xiao, Y., Lai, Y., R., and, Plaxco, W., K. Preparation of electrode-immobilized, redox-modified oligonucleotides for electrochemical DNA and aptamer-based sensing. *Nature Protocols*. 2007;2; - 2875 - 2880. doi 10.1038/nprot.2007.413