The G-quadruplex-hemin DNAzyme can tag itself with the reactive substrate biotin tyramide

by

Owen Einarson

B.Sc., Simon Fraser University 2013

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Chemistry
Faculty of Science

SIMON FRASER UNIVERSITY
Spring 2017

Approval

Name: **Owen Josef Einarson**

Degree: **Master of Science (Chemistry)**

Title: ***The G-quadruplex-hemin DNAzyme can tag itself with the reactive substrate biotin tyramide***

Chair: Robert A. Britton
Professor

Examining Committee:

Dr. Dipankar Sen
Senior Supervisor
Professor

Dr. Jeffrey J. Warren
Supervisor
Assistant Professor

Dr. Peter J. Unrau
Supervisor
Professor
Department of Molecular Biology and Biochemistry

Dr. Lisa Craig
Internal Examiner
Professor
Department of Molecular Biology and Biochemistry

Date Defended/Approved: April 19, 2017

Abstract

Genomic database searches suggest that there are a large number of potentially G-quadruplex forming sequences present in the human genome, in addition to their well-established localization in telomeres. Novel imaging techniques support these data and have begun to indicate the extent to which G-quadruplexes are present *in vivo*. Information on any biological function of these sequences is less clear. This thesis presents the creation of an assay that could target G-quadruplexes for imaging or pulldown and potentially give information on their biological relevance. We utilize the inherent peroxidase activity of the quadruplex-hemin interaction to initiate a tagging reaction based upon a reactive tyramide substrate.

Keywords: G-quadruplex; peroxidase; DNAzyme; aptamer; tyramide;

Acknowledgements

I owe a great deal of thanks to Dr. Dipankar Sen for his continued patience and support during my studies. He kept me on track through demoralizing moments and has helped me learn a lot about the process of scientific research.

Thank you to my committee: Dr. Peter Unrau, and Dr. Jeff Warren, for your very helpful feedback and constructive criticisms during committee meetings.

To all the Sen Lab members and Unrau lab members who have helped along the way, and given me a great support system, thank you.

Table of Contents

List of Figures

List of Acronyms

Å	angstrom
APEX2	Modified ascorbate peroxidase
ATP	Adenosine triphosphate
bp	Base pair
EDTA	Ethylenediaminetetraacetic acid
G4	CatG4 parallel G-quadruplex
NMM	N-methyl mesoporphyrin IX
nt	Nucleotide
PCR	Polymerase chain reaction
PDS	Pyridostatin
T_M	Melting temperature

Chapter 1.

Introduction

1.1. Overview

The research presented in this thesis covers the analysis of different functional DNA structures. These structures are all designed, or selected, through explicitly non-biological means, and studied *in vitro*, though some of the analyses will attempt to connect the *in vitro* findings with biologically relevant situations. Chapter 2 briefly describes data relating to a designed DNA switch, that can be controlled by the presence of the mercuric ion. This research is linked loosely to the following two chapters, but serves as an example of the non-biological applications that DNA can fulfill. Chapter 3 will change focus to analysis of the structure and hemin binding characteristics of DNA oligonucleotides possessing 2 stranded duplex and 4 stranded quadruplex regions. These designed DNA sequences may have nanotechnology applications and potential ramifications for the discovery of similar structures in nature. Chapter 4 concentrates on a reaction arising from the quadruplex-hemin interaction that may be beneficial for tracking and identifying G-quadruplex folded DNA in living cells.

This chapter introduces relevant background information relating to the structural variability of DNA. Specifically, the structural and biological implications of alternative base pairing, as well as how DNA is able to form four stranded quadruplexes. It also introduces the relevance of hemin-nucleic acid interactions and the potential importance of quadruplex-hemin interactions *in vitro* as well as *in vivo*.

1.2. DNA in biology

DNA is most often associated with its role in the storage and propagation of genetic information. This information is transcribed into RNA and translated into the proteins that are the primary players of biological machinery. The relative structural and chemical simplicity of the DNA macromolecule, compared to complex protein molecules, leads to less opportunity for it to take part in the many chemical reactions occurring in biological

chemistry. This same simplicity makes it well suited for the stable preservation and accurate propagation of an organism's genes.

As we continue to refine our understanding of the biological and synthetic forms of nucleic acids, the direct relationship between sequence and structure has proven useful in designing novel and functional DNA structures for purposes of genetic engineering, drug design, and nanotechnology. The inherent stability of DNA allows it to fulfill an increasing number of roles as a research tool.

It is now well understood that DNA and RNA perform many other functions alongside their primary role in carrying genetic information. DNA elements that do not function in coding for protein may act in ways with direct or indirect impact on cellular functions. DNA codes for different functional—as opposed to protein-coding—RNA molecules like ribosomal RNA, transfer RNA, and small interfering RNA. A gene's level of transcription can be up- or down-regulated by non-coding DNA elements that bind to transcription factors. Telomeres are large repetitive non-coding regions at the ends of chromosomes, which buffer the genome from degradation. With less than 2% of the human genome coding for proteins,[1] an understanding of the role of much noncoding DNA is still being sought.

1.3. DNA structure

The basic unit of DNA is the nucleotide, made up of a phosphate group linked to a 5-carbon sugar which is linked to one of several possible nitrogenous bases. The vast majority of DNA is composed of 4 nucleotides, containing the bases adenine, guanine, thymine and cytosine. Adenine and guanine are purines, containing fused 5 and 6 membered heterocycles. Thymine and cytosine are composed of only a 6-membered heterocycle and are called pyrimidines. The complementarity of hydrogen bond donors and acceptors on each of these bases gives rise to specific base pairing schemes: adenine (A) will generally pair with thymine (T) and guanine (G) with cytosine (C). When DNA is synthesized, or enzymatically replicated from a parent molecule, the result is a strand of nucleotides, linking sugar to sugar through phosphoester linkages through the phosphates, and the nucleobases linked to sugar via nitrogen to carbon N-glysosidic bonds. If a strand of DNA with the complementary base sequence to a reference strand is present (as it would be in biological systems), the two strands will spontaneously interact

via base complementarity to form a double helix structure. In the double helix, the complementary strands lie in a mutually antiparallel orientation. This antiparallel macromolecular complex is known as a DNA duplex. The bases on opposite helices pair by orienting themselves to the center of the helix. In the B-family of double helices favoured by DNA (as opposed to the A-family favoured by RNA double helices) the base pair heterocycles are coplanar, with the plane lying almost perpendicular to the helix axis. The stability of a duplex arises from noncovalent interactions of two general types.[2] The first is the favoured enthalpy of hydrogen bonding between complementary bases in the two strands; this is favourable enough to overcome the loss of hydrogen bonding with the solvent and entropic effects. Another important factor contributing to duplex stability is base stacking, which occurs to a greater extent in double stranded DNA versus single strands. Base stacking arises from favorable overlapping of pi orbitals in the aromatic base heterocycles, and from beneficial transient and permanent electrostatic interactions as well as hydrophobic interactions. The planar orientation of base pairs above and below one another makes this possible, even though the actual overlap of base pairs is slightly offset in the helix.

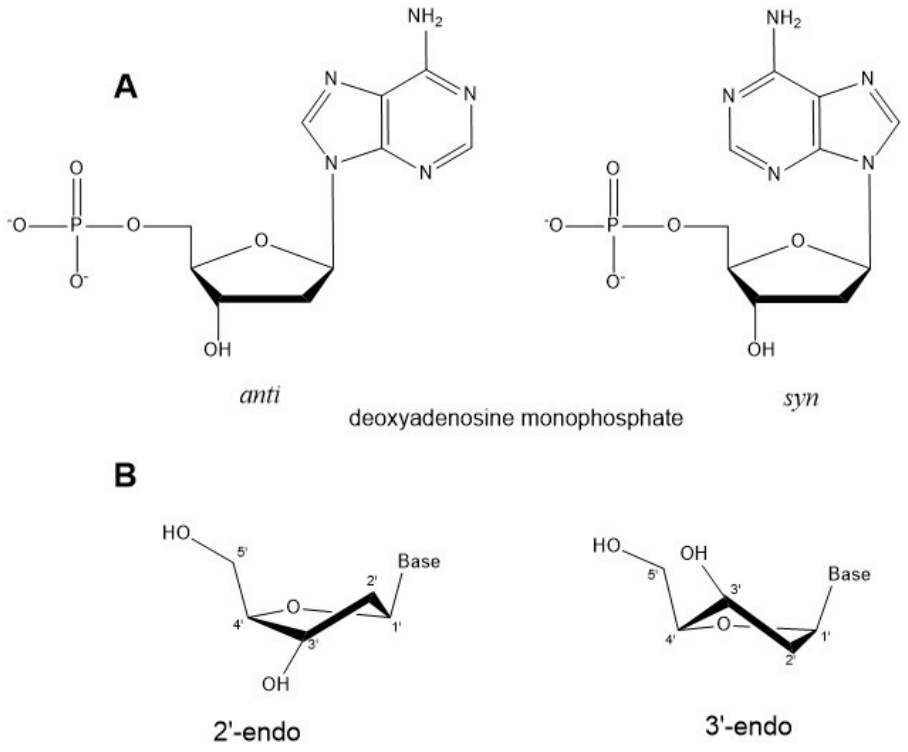


Figure 1-1 Nucleotide glycosidic orientation (a) and sugar pucker (b).

1.4. DNA structural variability

The most common structure adopted by duplex DNA is known as a B-type helix, which is a right-handed helix with ~10.5 base pairs per full turn. This arrangement represents the overwhelming majority of DNA found in a cell but there are a number of other biological and non-biological conformations that can be induced under certain conditions. Z-type and A-type DNA are examples of other duplexes present in biology. Z-type is a left-handed helix that forms only transiently under specific conditions, such as a strictly alternating repeating GC sequence or torsional strain. Z-DNA constitutes a paradigm of some of the "unusual" structural features that are manifested in non-B helices. One is "glycosidic orientation," alternative orientations about the N-glycosidic bond of the base with respect to the sugar. Another is "sugar pucker" involving orientation of carbons in the sugar above or below the plane of the ring (Figure 1-1).

The bond between the base heterocycle and the sugar is the N-glycosidic bond, which can rotate to give two stable conformations: anti and syn. Anti is favoured owing to the lower steric hindrance between the atoms of the base and the sugar. This is the glycosidic orientation that is present in B-type DNA. In Z-type, the bases alternate between syn (for pyrimidine nucleotides) and anti (for purine nucleotides).

The planarity of the sugar also takes part in the outcome of DNA structure and two types of "sugar pucker" are common (Figure 1-1b). In the case of 2'-endo, present in B-type DNA, the oxygen molecules attached at the 3' and 5' positions are further apart than in 3'-endo, where they orient to the same face off of the sugar. This has the effect of narrowing the phosphate to phosphate distance in the sugar phosphate DNA backbone, allowing for different angling of the backbone. In Z-type DNA the sugar pucker alternates between the two situations; purines are 3'-endo and pyrimidines are 2'-endo.

In B DNA A-T and G-C pairs hydrogen bond in a specific pattern, called Watson-Crick base pairing (Figure 1-2). An A-T pair forms 2 H-bonds and a G-C forms 3 H-bonds, leading to G-C rich sequences having a greater resistance to unwinding. Another feature of Watson-Crick base pairing is the asymmetrical arrangement of the sugar-phosphate backbone. As the bases pair, the backbones orient such that they are closer on one side, leading to the presence of an imbalanced double helix. One groove of the helix is wider, and is known as the major groove. The smaller side of the helix is called the minor groove.

The dimensional difference between the major and minor grooves is large in B-DNA, smaller in A-DNA, and characteristically different from either in Z-DNA.

Watson-Crick base pairing is not the only way that the nucleobases, A, G, C, and T can pair with eachother or with themselves. Under different circumstances, other base pairing arrangements may occur. Two examples out of many are the Reverse Watson-Crick and Hoogsteen base pairing schemes. Reverse Watson-Crick occurs when one base is flipped 180° with respect to Watson Crick, but uses the same binding face. This pairing mode is found often in RNA, and so is shown in the figure with U (uracil) instead of T. The more relevant type of alternative base pairing for this thesis is the Hoogsteen base pair, which uses a different binding face for one of the nucleotides involved. In the case of a G-G Hoogsteen base pair, the C6-N7 face of one G hydrogen bonds with the Watson Crick face of the other G. Sugar pucker, hydrogen bonding on alternative faces of a base, and rotation about the sugar-to-base N-glycosidic bond open up many avenues for structural diversity.

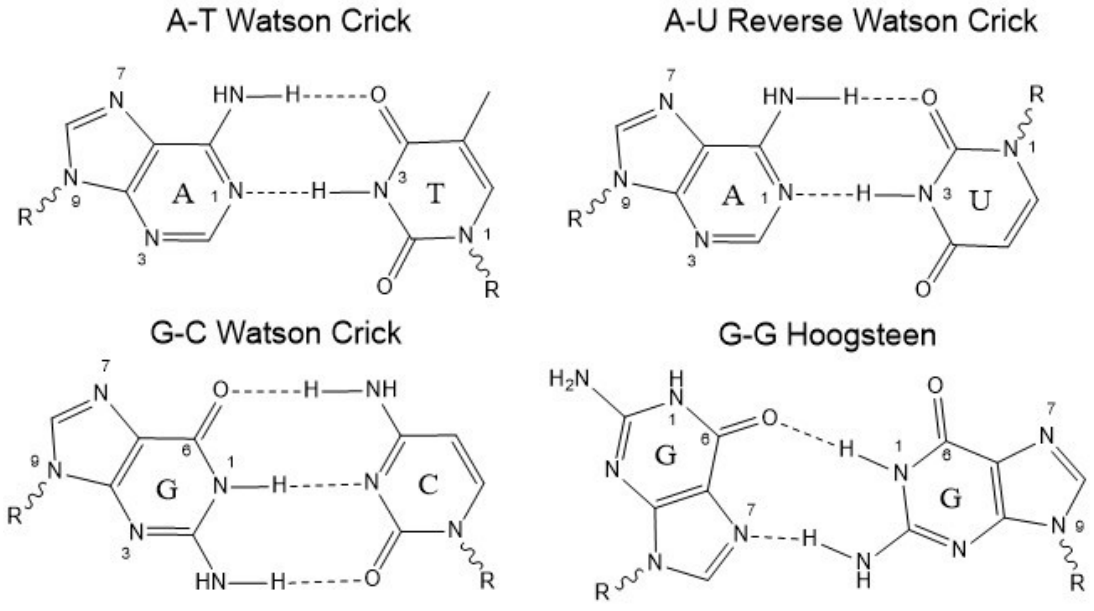


Figure 1-2 Nucleic acid base pairing is most often Watson-Crick in duplex DNA. Non-Watson-Crick base pairing can stabilize and allow for more exotic structures.

An important distinction to be made between DNA and RNA is the lack of the 2'OH group in DNA (Figure 1-1). One major chemical implication for the presence of this group is susceptibility of RNA to alkaline hydrolysis, and it is likely a key underlying reason why DNA is used as the long-term receptacle of genetic information in contemporary

organisms. The presence of the 2'OH group also has broad implications for structure and function, and its absence generally underlies why RNA has more diverse biological roles. Evolution has favoured biological RNA as a generally single stranded molecule. The resulting lack of base pairing leads to scenarios where the bases are free to associate with proteins and other molecules or to enable the strand of RNA to fold back on itself and form intra-strand base pairs. Compositely, RNA may form complex secondary, tertiary and quaternary structures, such as the well-known functional RNAs: transfer RNA and ribosomal RNA. The presence of the 2'OH translates to RNA preferring an A-type double-helix, with a wider minor groove opened up for molecular interactions. The combination of the above two factors means RNA is able to form varied and stable interactions with proteins. RNA structures that show the ability to carry out enzymatic processes, often aided by the 2'OH functionality, are known as ribozymes. Ribozymes were first discovered in 1982. Thomas Cech discovered a pre-ribosomal RNA in *Tetrahymena* that could splice out its own introns in the absence of participation by any proteins.[3] At the same time, Sydney Altman discovered a bacterial ribozyme, RNAase P, capable of processing pre-transfer RNA into its mature, active form.[4] These discoveries led to an enlarged view of the possible catalytic roles of RNA and spurred on further investigation and discovery of many more enzymatically active RNA molecules.

DNA, however, has still largely been viewed as an inert polymer. While it is possible for DNA in its basic B-type structure to interact with proteins, the two-stranded base pairing arrangement in biology largely protects the information carrying bases from potential damage. The negatively charged phosphates on the DNA backbone encourages non-specific ionic bond formation with proteins (e.g., histones in eukaryotic organisms). In doing so, histones are able to regulate the compaction and availability of DNA for transcription. Sequence specific DNA-protein interactions began to be discovered in the form of transcription factors. These proteins most commonly insert DNA-binding peptide domains into the major groove of DNA, allowing amino acid side-chains to recognize specific base sequences. As an example, the basic leucine zipper domain (bZIP domain), present in many eukaryotic regulatory proteins, is specific for ACGT sequences in DNA.[5] Basic lysine and arginine residues interact with the major groove of DNA in a sequence specific manner. When present in the promoter region of a gene, this type of interaction can stimulate gene expression by binding to a transcription factor. Many types of these transcription factors that are capable of either up- or down-regulating gene expression

have been identified.[6] Likewise, transcription factors can serve to recruit other proteins to the DNA region for the same purpose.[7] These types of interactions are distinct from the capabilities of RNA to maintain functional structures by themselves. However, it was in the early 1990's when the study of DNA as a more functional molecule began to take shape through several key discoveries.

1.5. Selection of DNA aptamers

One of key technical innovations in the field of DNA and its structural and functional studies was the development of an *in vitro* technique dubbed 'systematic evolution of ligands by exponential enrichment' or SELEX. This technique, pioneered simultaneously by the Gold[8] and Szostak[9] groups, involved the creation of large, random-sequence pools of single stranded nucleic acid sequences (10^{14} to 10^{15} different sequences) and search this pool, in an iterative fashion, for individual DNA sequences capable of binding to a designated ligand. Initially, the two groups used SELEX for RNA, as it was known that RNA is capable of forming complex folds and structures. However, subsequent research using DNA found very little difference in affinity for DNA or RNA sequences selected against the same ligand,[10] with the added benefit of DNA being less susceptible to chemical degradation.

The key aspects of SELEX are the ability to separate sequences that have some binding affinity for the ligand from those that do not. This is generally done by linking the ligand to a column or bead, applying the pool, and washing away the unbound sequences. The resulting pool is enriched for ligand binders. The new pool is then duplicated and amplified through PCR and reapplied to the ligand column, thus further enriching true binders. Through multiple iterations of this process the enriched pool may come to exhibit a strong binding affinity for the ligand. Such a pool can then be sequenced and studied in depth to identify the region of the nucleic acid responsible for ligand interaction. The active portion of the nucleic acid is known as an "aptamer".

RNA and DNA aptamers have now been identified for a wide variety of molecules. Aptamers have been derived for whole cancer cells, recognizing differences in cell surface proteins, for the purposes of imaging a drug targeting.[11] Aptamers for many proteins and small biomolecules such as ATP,[12] glucose,[13] and bisphenol A[14] have also been reported.

With such a wide range of demonstrated targets aptamers offer an alternative to antibodies in many clinical and research applications.

1.6. DNAzymes

Another important discovery occurred in 1994, when the first example of a catalytic DNA was found.[15] This synthetic "DNAzyme" was capable of cleaving a nucleic acid sequence at the 3'-O-P bond of a ribonucleotide placed within a single-stranded, all DNA sequence. It was obtained using a SELEX in the presence of Pb^{2+} ions as potential hydrolytic cofactor, and the observed rate of cleavage was ~100 fold higher than that of the uncatalyzed reaction. The substrate was 19 nucleotides long and the DNAzyme contained 38 nucleotides. The substrate and DNAzyme base paired at their 5' and 3' ends, with the unpaired central regions and likely coordinating the Pb^{2+} to catalyze the bond cleavage. More variants of RNA and DNA cleaving DNAzymes were found, with the RNA variants reaching kinetic efficiencies approaching the rate of diffusion ($10^9\ M^{-1}\bullet sec^{-1}$).[16–18] The range of enzymatic functions that has been demonstrated for DNAzymes to date is narrower than for ribozymes, but examples of catalytic DNA have grown to include ligases,[19] C-C bond formation,[20] thymine dimer photorepair,[21] and most importantly for this thesis, peroxidases.[22] In contrast to the various naturally occurring ribozymes, no naturally occurring DNAzymes have yet been isolated.

Work by Li *et al.* in 1996 attempted to expand upon the scope of DNAzyme functionality by selecting for a DNA aptamer that bound to N-methylmesoporphyrin (NMM) with the intention of finding a DNA catalyst for metal insertion into a porphyrin[23]. NMM is a known inhibitor and stable transition-state analogue of cellular ferrochelatase enzymes. It was known that when antibodies are raised against NMM, they also gain the ability to act analogously to natural ferrochelatase, and catalyze the insertion of metals like Mn^{2+} and Zn^{2+} into mesoporphyrin IX (MPIX)[24]. The structure of NMM mimics the transition state of mesoporphyrin IX during metal insertion by ferrochelatase, by virtue of it being less planar than the underivatized mesoporphyrin IX molecule.

The successful SELEX experiment of Li et al. resulted in a number of G-rich single-stranded DNA sequences with a high affinity for binding NMM.[23] Two of these G-rich variants containing 24 (PS5.M) and 21 (PS2.M) nucleotides were then shown to be DNAzymes capable of inserting Cu^{2+} and Zn^{2+} into MPIX. A requirement of potassium for

catalysis and the guanine richness of these DNA sequences suggested that they folded to form a structure known as a G-quadruplex.

1.7. G-quadruplexes

The ability of guanines to form alternate structures was first identified by melting studies done on short, G-G-G oligonucleotides in 1962. In these experiments aggregates would begin to form that could be dispersed through heating.[25] Soon after, mixtures of guanylic acid that formed viscous gels were analyzed by X-ray diffraction and a model was proposed for a base-pairing scheme called a guanine-quartet (G-quartet)[26] (Figure 1-3). It was also correctly proposed that two or more of these structures should stack well into, owing to the large planar surface of a G-quartet. Stacked guanine quartets have come to be know as G-quadruplexes.

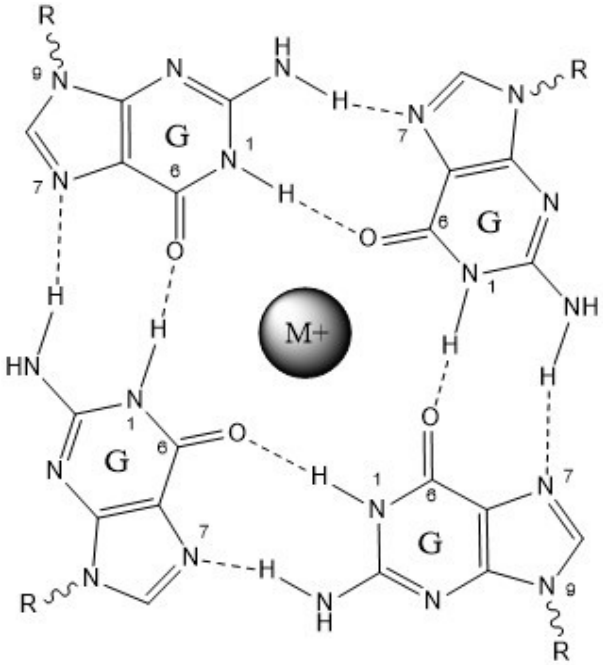


Figure 1-3 Hoogsteen base pairing in a guanine quartet. The central metal ion (M^+) stabilizes the structure by coordinating the keto- oxygen atoms. When G-quartets form multiple layers it is known as a G-quadruplex.

In 1988, Sen and Gilbert found DNA G-quadruplexes occurring *in vitro* in oligomers based on the guanine-rich immunoglobulin switch region genes.[27] They found that these sequences began to interact and aggregate in a concentration dependant manner after being stored in the dilute solution of a sodium salt. Gel mobility experiments showed low

mobility bands, representative of the formation of an interstranded tetramer structure. Attempted methylation of the complex with dimethyl sulfate (DMS) showed that guanines in the complex were protected. DMS will methylate guanine at the N7 position in single-stranded or double-stranded DNA. Protection in these multi-stranded structures supported the presence of Hoogsteen base pairing that involves the guanine N7 positions. The model proposed to account for the tetramer gel pattern and methylation pattern is represented in Figure 1-4a, where the strands align in a parallel direction. Later experiments would show the important role that salt, different cations, and sequence play in determining the structure, as well as overall topology of G-quadruplexes.[28] Stabilizing cations such as Na^+, K^+, and Sr^{2+} coordinate to the keto-oxygen atoms in the central space of the quartet, and the extent of coordination is largely determined by the size of the ion.[29] Lithium's small size coordinates poorly, whereas more physiologically relevant Na^+ and K^+ coordinate and stabilize G-quadruplexes. Na^+ binds in the plane of the quartet, and K^+ sandwiches between two quartets. It was found that changing the ratio of Na^+/K^+ in the folding buffer of G-rich telomeric sequences influenced whether these oligonucleotides would fold to tetramolecular or bimolecular quadruplexes.[28] Low K^+/high Na^+ enhanced the kinetics of tetramolecular G4 formation versus Na^+ alone. Higher K^+ would only yield a bimolecular quadruplex structure, where a strand folded back on itself in antiparallel G-G base pairing, then dimerized to form a quadruplex. Separate experiments also probed telomeric sequences using the UV-crosslinking of thymines. These experiments showed that single strands with appropriate spacing of guanines, i.e. separated runs of guanines, could fold back on themselves and form intramolecular quadruplex structures in an antiparallel arrangement.[30] These findings highlight the different possible quadruplex morphologies and many more since then have been shown to be possible (See review[31]). Figure 1-4 illustrates 3 conformations of intramolecular quadruplexes with parallel(d), antiparallel(b) or mixed orientations(c).

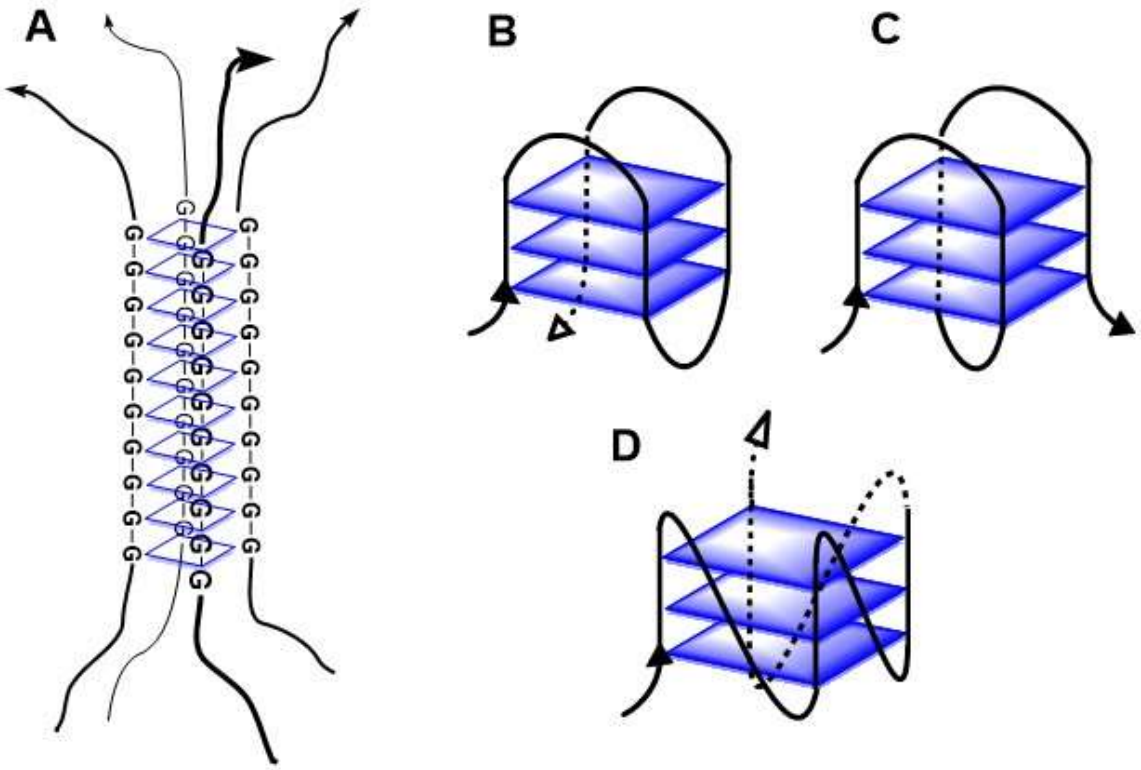


Figure 1-4 G-quadruplexes can be intermolecular (a), or intramolecular(b-d), and the directionality of the strands in the quadruplex can be parallel (a, d), antiparallel (c), or mixed (b).

Since the discovery that telomeric DNA can form quadruplexes *in vitro,* studies have identified their formation *in vivo,*[32,33] and functional implications also have been suggested.[34–36] Due to the relative abundance of telomeric DNA in cells, they have often been the focus when it comes to G-quadruplex function *in vivo*. Telomeric quadruplexes appear to both inhibit the telomere elongation by telomerase[37] and prevent telomere degradation by nucleases.[34] However, it has been more difficult to demonstrate the association of G-quadruplexes with *in vivo* functionality in regions that are less G-rich and where they may form transiently.

Successful visualization of quadruplexes *in vivo* shows their cell cycle dependent formation in the telomeres of ciliates. This assertion was made possible using cell staining with fluorescently labelled antibodies against specific quadruplexes.[38] Mammalian cells have now also been imaged with a similar antibody based approach.[33] The antibody was found to bind strongly to different conformations of G-quadruplexes and with similar affinity. Staining of the cell nuclei indicated the presence of quadruplex structures. Upon investigation of antibody binding to individual chromosomes during metaphase, when

chromosomes adopt the most condensed and visible form, it was found that the antibody bound at many sites. Among 100 sample chromosomes 58 demonstrated antibody binding. Of the binding observed in the 58 chromosomes, 75% appeared in non telomeric regions. This result suggests that stable quadruplexes may be widely distributed outside telomeres which is a very important observation.

1.8. Heme-quadruplex interactions

During the porphyrin metalation experiments, it was found that hemin (Fe^{3+} heme)—note that hemin and mesoporphyrin have very similar strutures—could act as an effective competitive inhibitor to mesoporphyrin IX metalation by PS5.M and PS2.M.[23] Heme is a ubiquitous molecule in biology carrying out a wide variety of catalytic and other functions. These functions include electron transfer, oxygen transport, and oxidation. Common features in most common heme enzymes is axial coordination of the heme iron to an amino acid side chain ligand on the protein.

A comparison of the visible spectrum of the PS2.M-hemin complex to heme enzymes, like the oxidized form of myoglobin, showed a great degree of similarity.[39] The possibility of the quadruplex-hemin interaction providing a similar functionality was confirmed by peroxidation experiments. Upon addition of the oxidant H_2O_2 to a mixture of PS2.M and hemin, the rate of peroxidation of a chromogenic substrate was enhanced by 250 times over a mixture of hemin alone.[39] Changes in spectral characteristics of hemin in the presence of G-quadruplex offer some insight into the interaction. Enhanced absorbance in the Soret region (~400 nm) of hemin is associated with an increase in the hydrophobicity at the binding location, and is seen in both heme enzyme[40] and hemin-quadruplex interactions.[39] The iron (III) moiety in oxidized myoglobin has six-fold coordination, with the axial positions coordinated to histidine and superoxide. The case is likely the same for the hemin-quadruplex complex, where the 6^{th} axial position is coordinated with water and is available to exchange with peroxide under the correct conditions. The other axial position is thought to be coordinated with a guanine. Footprinting experiments and molecular docking simulations indicate that hemin binds by end stacking upon a G-quartet.[41,42] EPR spectroscopy experiments support this model, and also identified the presence of a carbon-centered radical forming on the DNAzyme during peroxidation.[43]

The PS2.M-hemin peroxidation reaction was determined to be pH sensitive and was maximal near pH 8.0.[39] This pH sensitivity is linked with the “alkaline transition" of the axial water ligand; deprotonation of the heme water ligand leads to a slower ligand exchange with peroxide and a slower reaction[44]. The ability of the enzyme to provide an environment where deprotonation at near neutral pH is less likely is important for peroxidation. Uncomplexed hemin shows a pK_a of ~4, while horseradish peroxidase has a pK_a of 10.9. With a pK_a of 8.7, PS2.M-hemin is able to function efficiently at near neutral conditions. The peroxidation reaction is also enhanced by nitrogenous buffers, such as ammonium, which may act in acid and base roles during the activation of peroxide, as distal residues do in heme proteins.[39]

The PS2.M-hemin complex can oxidize a broad range of substrates. PS2.M-hemin exhibited a 70-fold rate enhancement versus hemin alone, and performed at faster rates than horseradish peroxidase for a number of phenolic substrates.[45] Active sites in peroxidase enzymes confer enantioselectivity when oxidizing certain substrates. PS2.M-hemin lacks enantioselectivity for any substrates, likely due to the less constrained binding site.[45]

Recent work in our lab by Shumayrikh *et al.* investigated whether or not other multi-stranded nucleic acid structures (isoguanine quadruplexes and pentaplexes) were also capable of hemin induced peroxidase function.[46] While hemin and isoguanine (iG) pentaplexes exhibited peroxidase activity similar to that of G-quadruplexes, iG quadruplexes did not. This was attributed to the known non-planarity of the iG quadruplex quartet. G-quadruplexes and iG-pentaplexes are both planar, suggesting that a distally coordinated nucleobase and planarity may be sufficient hemin activation.

Chapter 3 of this thesis further analyzes the specificity of hemin-quadruplex interactions by investigating quadruplex-duplex hybrid structures containing unique binding environments. Chapter 4 then investigates the peroxidase activity of the hemin-quadruplex interaction as a means for self-tagging by way of a reactive biotinylated substrate. However, Chapter 2 will first briefly diverge into research unrelated to hemin-quadruplex interactions, and investigate a designed DNA switch that responds to Hg^{2+}.

Chapter 2.

An aptamer-based switchable Hg^{2+} sensor

2.1. Introduction

The ability of DNA and RNA to tightly bind target molecules with high affinity make them useful candidates for use as biosensors. Commonly, the addition of a redox label to the aptamer serves as the mechanism to translate the physical change upon ligand binding into an electrochemical signal. Many designs rely on an aptamer attached to a redox label, tethered to an electrode, with the distance from electrode to redox label directly determining the rate of electron transfer.[47,48] Different aptamer designs have been implemented to allow for the most effective sensing: single stranded, electrode bound DNA; double stranded DNA with the aptamer binding ligand displacing one strand. A promising technique used by the Sen and Yu labs in recent years has been to utilize a duplex structure, where the aptamer controls the ability of the duplex to form.[49] The efficiency of charge transport is dependent upon the integrity of the duplex instead of distance from redox label to electrode surface. For the aptamer-analyte binding event to bring about conformational change sufficient to affect duplex stacking, it must be intimately linked or embedded within the duplex. Both of these designs were successfully applied to the sensing of an adenosine target.[50]

The work in this thesis was based upon the design of a switch where the aptamer is not directly part of the duplex. The linkage responsible for transposing the conformational change in the aptamer to the duplex was structured around a 3-way junction. By separating two conductive, duplex regions with an aptamer binding domain, the extent of interhelical stacking can be modulated by the addition or subtraction of analyte. This method was applied successfully to the sensing of both thrombin and lung cancer biomarker CTAPIII, where the presence of these analytes caused a conformational change allowing for an increased conductivity and therefore an increase in the electrochemical signal in a concentration dependent manner.[51]

Thomas *et al.* more recently implemented this technique to detect mercuric ion.[52] Hg^{2+} is a highly toxic pollutant found as a by-product in many industrial applications. It can lead to long term health problems such as brain damage. Current methods to detect Hg^{2+}

in the environment rely on atomic absorption spectroscopy, inductively coupled plasma mass spectrometry, and selective cold vapor atomic fluorescence spectrometry, which can involve long preparation procedures and expensive equipment. More recent methods applying colorimetric methods can lack sensitivity. As an alternative, Thomas *et al.* utilized the capability of thymine-thymine base pair mismatches to be stabilized by a bridging Hg^{2+} ion, forming a stable T-Hg^{2+}-T pair.[53] The analyte binding region of the sensor was designed with multiple T-T mismatches, and upon addition of Hg^{2+} they were able to measure an increase in conductivity. The most successful design was a variant called 3'AA. However, it was noted that several of the designed variations showed properties that differed from the standard condition where increasing the analyte concentration gave rise to an increasing signal. This was most evident in a variant called NB-1. The goal of the research presented here to further investigate the properties NB-1, which showed a reduction in signal upon Hg^{2+} addition.

2.2. Materials and methods

Oligonucleotide preparation. Oligonucleotides were synthesized by Integrated DNA Technologies (IDT). One oligo (C1) was purchased with a 5' modified with a $H_2N(CH_2)_6$ group. Oligonucleotides were treated with 10% aqueous piperidine and heated to 90ºC for 30 minutes to cleave any pre-existing sites of DNA damage. Samples were then ethanol precipitated, purified by denaturing PAGE, eluted by crushing the gel into small pieces and soaking in TE buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0), ethanol precipitated, and resuspended into ddH_2O. Anthraquinone (AQ) conjugates were prepared by reaction of 5'-C6-NH_2-modified oligos (C1) with the NHS-ester of anthraquinone-2-carboxylic acid as described previously[50]. AQ-modified oligos were purified by HPLC through an Agilent Eclipse XDB C-18 3.5um 4.6x100mm column; solvent A: 0.1 M triethylammonium acetate (pH 7.0)/CH_3CN (92:8) and solvent B: CH_3CN; 30 min. linear gradient from 0% B to 40% B; 1 mL/min. Hg^{2+}-binding oligos were 5'-^{32}P-labeled using OptiKinase (Affymetrix) and γ-^{32}P-ATP (Perkin-Elmer) following manufacturer's instructions, and ethanol precipitated. Aqueous $Hg(OAc)_2$ stocks and dilutions were prepared fresh for each experiment.

Oligonucleotide sequences

C1: 5′-$H_2N(CH_2)_6OPO_3$-TTT AGC TCA CGA GAC GCT CCC ATA GTG A -3′
NB-1*:* 5′-TCA CTA TGG GAG CGT TTT GTC GGG AGA CTT TTT CTC GTG AGC TAA A -3′
3′AA*:* 5′-TCA CTA TGG GAG CGT TTG TTT GCG GGA GCT TTC TTA AAT CTC GTG AGC TAA A -3′

DNA charge transport experiments. Constructs were annealed by heating AQ-labeled oligos (AQ-C1) and the 5'-^{32}P-labeled Hg^{2+}-binding oligos to 95ºC for 2 min in 10 mM TrisOAc (pH 8.0), followed by cooling to room temperature over 30 minutes. At room temperature, binding buffer (5 mM $MgCl_2$, 50 mM TrisOAc, pH 8.0) was added with the desired concentration of $Hg(OAc)_2$. The AQ-C1 strand was at a concentration of 125 nM, and the Hg^{2+}-binding strand at 100 nM. Samples were transferred to a 96-well ELISA plate placed in an ice water bath, and irradiated in a 4ºC room. The lamp (UVP Black-Ray UVL56 365 nM) was positioned 4 cm above the plate for 30 min. Only the center 2 rows of the plate were utilized to maintain uniform irradiation. The samples were transferred to microcentrifuge tubes for ethanol precipitation. Resuspension in 100 µL 10% aqueous piperidine and heating to 90º C was followed by lyophilization, resuspension in 70 µL ddH_2O, and lyophilization again. The products were analyzed by 12% denaturing PAGE and autoradiography of the radioactively labeled DNA. Gels were placed beneath storage phosphor screens to capture ^{32}P exposure patterns. Autoradiography imaging and quantification of gels was performed on a Typhoon 9410 phosphoimager with ImageQuant TL7.0 software. Oxidative damage at reporter guanines was measured as the ratio of intensity (radioactive counts) of G_8 cleavage bands to the intensity of undamaged, full length oligos (ΔI). The [Hg^{2+}] dependence of G_8 damage was measured by plotting $\Delta I/I_o$ for all concentrations, where I_o is the fraction of G_8 cleavage observed in the absence of Hg^{2+}.

2.3. Results and Discussion

2.3.1. DNA biosensor charge transport in solution

Hole transport efficiency for 3'AA and NB1 is shown in Figure 2-1 through patterns of oxidative guanine damage. Samples were subjected to UV light treatment to induce the photosensitizer anthraquinone to inject an electron hole into one end of the DNA

switch. The extent to which the injected electron hole is able to migrate from the AQ attachment point, through the 3-way junction, and to the G8 guanine site to induce oxidative damage can be visualized directly by treating the samples with base to cleave the DNA at oxidized locations.

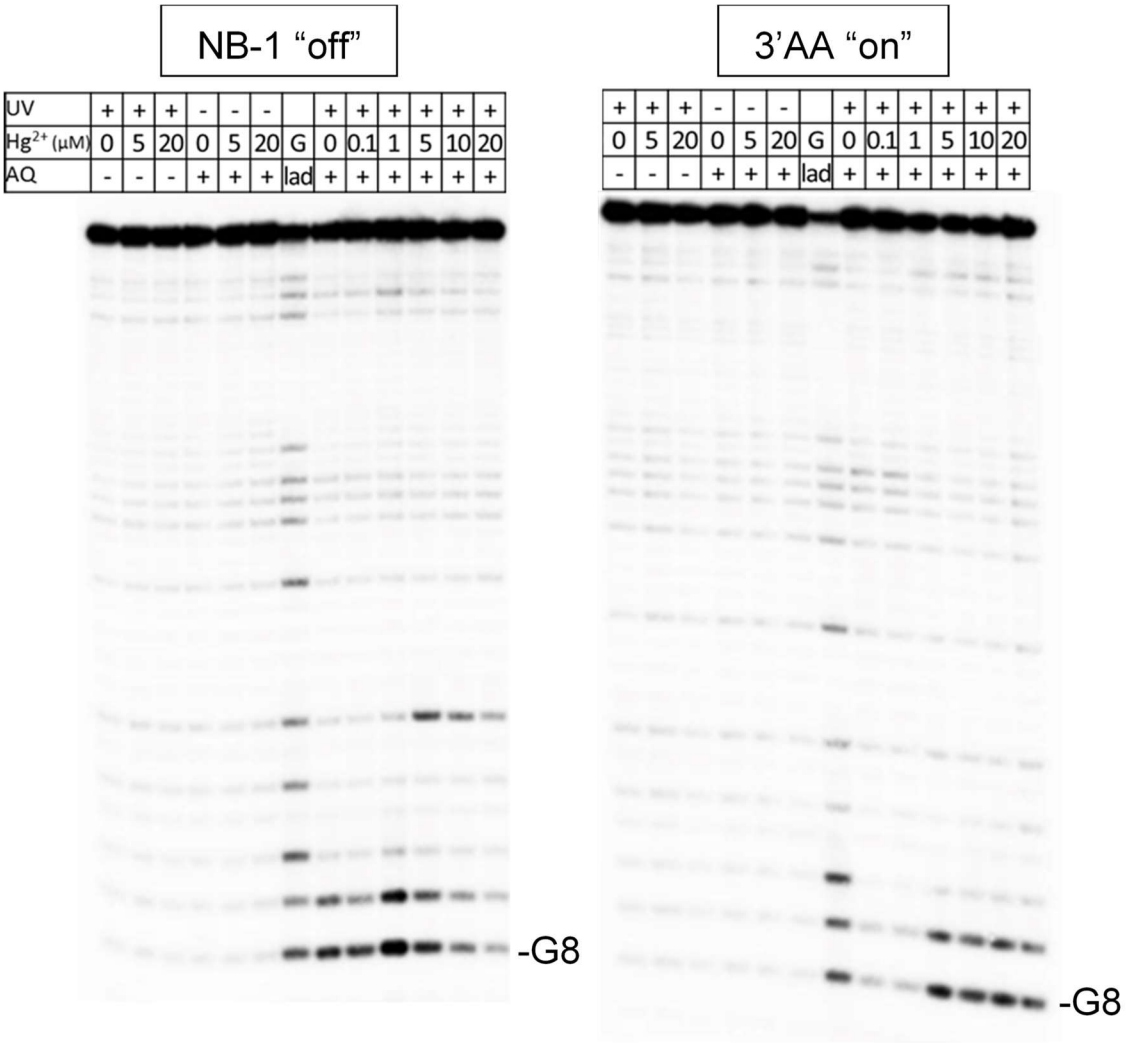


Figure 2-1 Oxidative guanine damage as a function of mercuric ion concentration. NB-1 shows an overall decrease in damage with increasing Hg^{2+}. The trend is reversed in 3'AA . Autoradiograph of 12% denaturing gel.

The construct 3'AA in Figure 2-1 represents the sensor capable of increasing signal in the presence of Hg^{2+}. On the right side of gel, the increasing Hg^{2+} concentrations increase the amount of oxidative damage at the distal G8 guanine site, indicating increased hole transport efficiency. This 3'AA data is a validation of previously reported

data.[52] The lanes with no G8 damage identify the absence of charge transport mediated oxidative damage when there is either no AQ or no UV light treatment.

The inverse effect is seen in the case of the NB1 construct (left side), where in the case of 0 μM Hg^{2+} there is already G8 damage and increasing Hg^{2+} leads to decreased charge transfer efficiency and damage. However, in NB1, the hole transport efficiency initially appears to increase, declining after 1 μM. This is not observed in the electrochemical data where these aptamers were tethered to gold electrodes (unpublished, not shown). Under those conditions, all additions of Hg^{2+} decrease the flow of current in a stepwise fashion. The range of Hg^{2+} induced signal change is greater in the 3'AA construct than the NB1 construct and can be seen in the magnitude of $\Delta I/ I_O$ values at 20 μM Hg^{2+} in Figure 2-2. This agrees with the greater signal change seen for 3'AA in the electrochemical data (not shown).

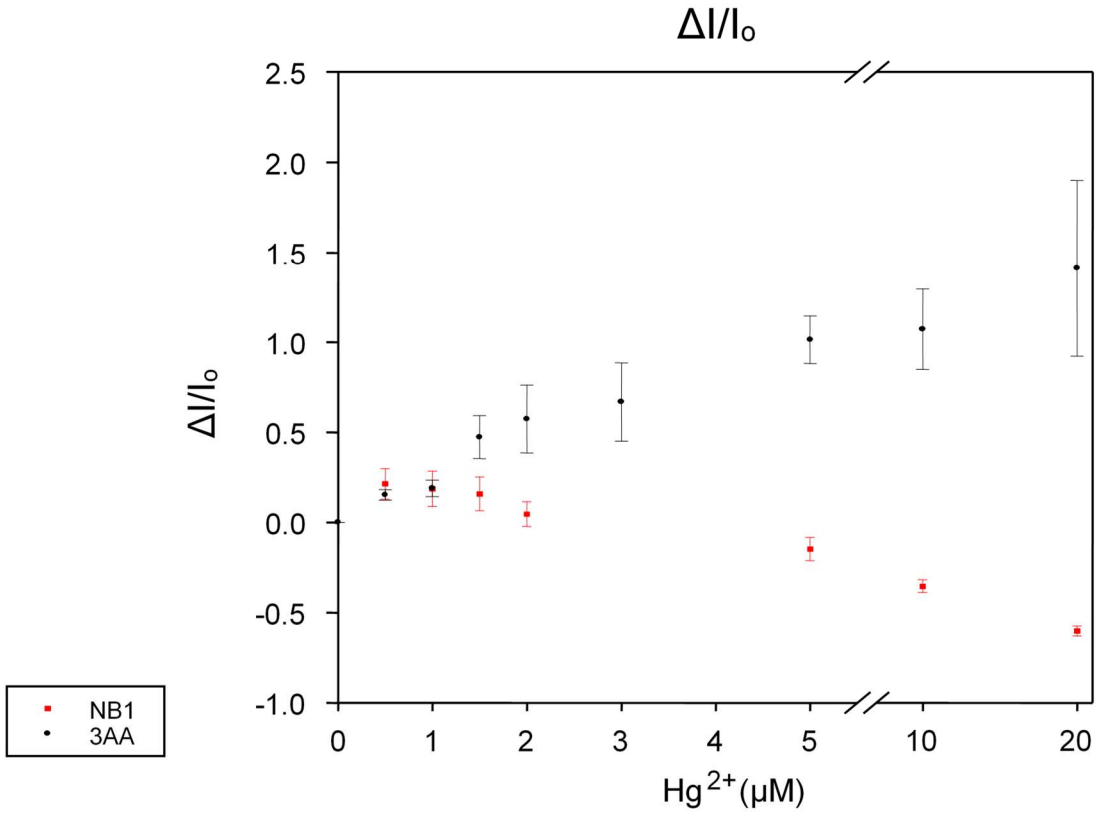


Figure 2-2 $\Delta I/ I_O$ for NB-1(red) and 3AA(black). Error bars indicate s.d. from 3 replicates.

Figure 2-2 shows the Hg^{2+} dependency of damage to G8 guanines represented as a ratio of G8 damage at a given concentration of Hg^{2+} (ΔI) versus G8 damage at zero Hg^{2+}

(I_O). Viewed this way, it can be seen that the “off to on” switch, 3’AA, continuously increases in damage with increasing Hg^{2+}. The “on to off” switch, NB1, begins above 0 before beginning to decline after ~2 μM Hg^{2+}.

2.4. Conclusion

The biochemical data here show these aptamer sensors have a low μM detection limit. However, this is not low enough to be useful in a Hg^{2+} detecting role. The electrochemical data from our collaborators in the Yu lab, using the same constructs on a gold electrode, showed a much better, low pM, detection limit. The reasons for this difference may come from the efficiency of electron transport in the electrochemical versus photochemical experiments.

Chapter 3.

Hemin binding properties of duplex quadruplex hybrid oligomers

3.1. Introduction

The ability of G-quadruplexes to bind hemin and catalyze peroxidative reactions has implications for biology, and potential industrial, and nanotechnology applications. A number of *in vitro* studies have observed G-quadruplex forming genomic sequences able to bind hemin.[54,55] Our lab has found that expansion of a $(G_4C_2)_n$ repeat within the human C9orf72 gene, previously associated with neurodegenerative diseases, binds and activates hemin.[56] It will be important for future work in this area to determine if any hemin-quadruplex activity exists *in vivo*. G-quadruplex-hemin complexes can also be valuable structural motifs for designed DNA nanotechnology. Due to their stability and the additional secondary structures they allow for the interaction has been used to create colorimetric biosensors[57] and switchable enzymes[58]. For these reasons, the structure and stability of different quadruplex-hemin motifs is valuable. The hemin binding characteristics of many of the most common quadruplex topologies have been investigated. Most orientations of quadruplex are able to bind hemin to some extent, but it is clear that parallel quadruplexes bind with the highest affinities and catalyze peroxidation most efficiently.[55] Antiparallel quadruplexes are capable of catalyzing effective peroxidation nearing levels of parallel structures in cases where the hemin has been covalently linked to the quadruplex DNA.[59]

The stability of quadruplexes allows for the design of a wide variety of more complex structures. Many of the non-guanine, non-base pairing regions of quadruplexes can be changed or extended without large penalties in melting temperature. There is much less information available about the hemin binding impact of extending quadruplex loops. In genomic DNA and RNA outside the telomeres, the opportunities for transient quadruplex formation is poorly understood.[60] The opportunities for quadruplex formation at times when DNA or RNA is single stranded (replication, transcription, translation) would be heavily impacted by the length of loops that could be accommodated. Recent work by Lim *et al.* has begun to explore the stability characteristics of a number of novel quadruplex folds.[61] The research presented in this chapter is intended to gain a greater understanding

of the hemin-binding characteristics of these unique G-quadruplexes, which may have implications for locating similar analogs in biology through the methods presented in Chapter 4.

3.2. Duplex quadruplex hybrids

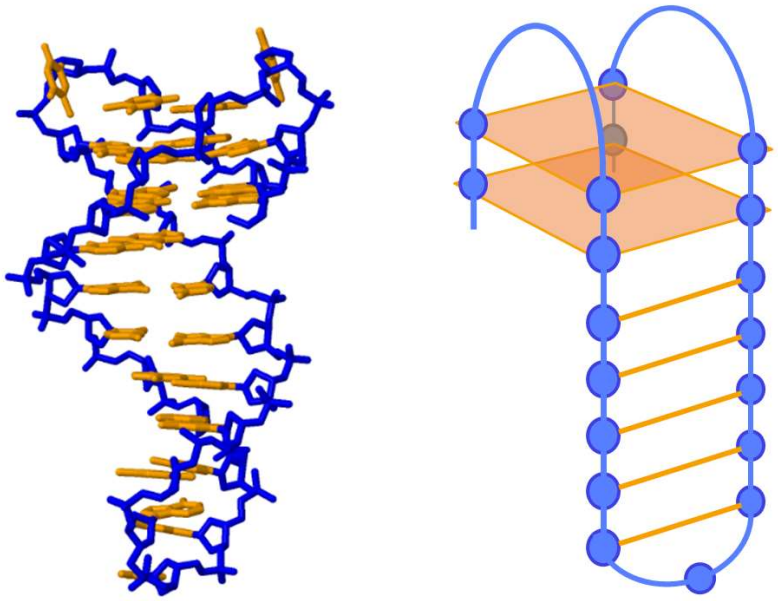


Figure 3-1 Quadruplex duplex hybrids (Construct 1 shown[62] [PDB 2M8Z]). NMR structure on left, cartoon representation on right.

The morphologies of quadruplex structures that have been previously described are generally identified in using algorithmic genome screens using strict rules with respect to loop length.[63] A common restriction is ≤7 nucleotides and Genomic screens of this type have yielded ~361 000 putative quadruplex forming sequences.[64] Other studies have found a correlation between loop length and the thermal stability of the quadruplex, with shorter loops having greater stability.[65] A necessity of any biologically relevant quadruplexes must be stability at physiological temperatures. However, analysis of extending loops in quadruplexes has focused on the addition of sequence that does not provide any stabilizing interaction with the core quadruplex. The work of Phan *et al.* explored the possibility of using stem loop structures to act as the loop regions of quadruplex structures in order to utilize possible stacking interactions between the tetrad and duplex regions with the intention of maintaining longer loop lengths without affecting melting temperature (T_m) adversely.[61] Stem loops occur when palindromic DNA folds back

on itself to base pair, incorporating unpaired loop bases at the point of reversal. They designed several constructs based upon anti-parallel, mixed orientation, and parallel quadruplexes. For this thesis, I will focus on 3 of these designs: Construct 1, Construct 3, and Construct 5.

Construct 1 (Figure 3-1) is a chair type antiparallel quadruplex, which has a duplex stem spanning the wide groove of the quartet. The wide groove spans 19 Å from phosphate to phosphate, and narrows to 18 Å at the duplex, creating a stable junction that allows continuous base stacking. This stability is demonstrated in melting experiments that showed the following properties for Construct 1 as a function of duplex loop length: T_m @ 0 base pairs=36.2 °C; T_m @ 3 base pairs=41.6 °C; T_m @ >6 base pairs≥44 °C.[61] Construct 3 (Figure 3-5), a basket type antiparallel quadruplex, attached a stem loop across diagonal corners of the quartet, meaning a phosphate to phosphate distance of >20 Å. To ease the transition of this distance to the 18 Å duplex portion, a purine-purine mismatch was introduced at the junction (no melting data). Construct 5 (Figure 3-6) was an all parallel G-quadruplex with 3 stacked G-quartets, with one of the propeller loops replaced with a stem loop structure. This structure required the breakage of a base pair at the junction for stability. The melting stability of this construct was much less dependent on duplex length; the initial quadruplex had a Tm @ 0 base pairs=62.2 °C, with minor fluctuations up and down with increases up to 16 bp. Unlike Construct 1, Construct 5 also withstood the loop being unmatched T's; at 35 T's the T_m was 41.6 °C. These characteristics can likely be explained both the presence of an extra quartet and by the orthogonal placement of the duplex in Construct 5. With the lack of stacking between the duplex and quadruplex, there is less influence of one region on another.

The stabilities of these novel quadruplex morphologies suggest that loop lengths in algorithms should be increased and that stem loops should be considered valid, or in some cases stabilizing. A recent bioinformatic screen of the human genome looking for quadruplexes containing stem loops (max. 20 nucleotides) identified ~48 000 putative sequences, ~33 000 of which were not found with the previously mentioned screen where the maximum loop was 7 nt.[63] Investigation of a number of selected sequences with NMR indicated many of these putative duplex-quadruplex hybrids contain Watson-Crick and G-tetrad regions. This evidence raises the possibility that current genomic quadruplex screening methods may not be inclusive for more exotic forms.

We wondered if the various loops would affect hemin binding and effectiveness in catalyzing oxidations. To investigate these questions, we selected Constructs 1, 3, and 5 as they appear to represent 3 distinct hybrid structures in terms of orientation and loop placement. They are compared with the parallel quadruplex CatG4; a quadruplex based upon PS2.M that has been modified for simplicity.

3.3. Materials and Methods

3.3.1. DNA oligonucleotides

All oligonucleotides were purchased from IDT. Samples were solubilized in a 10% piperidine and placed at 90°C for 30 minutes to cleave damaged oligos. After vacuum drying, the oligos were gel purified in 8-12% denaturing gels followed by ethanol precipitation and resuspension in TE buffer. For gel visualization, DNA was ^{32}P-labeled using OptiKinase (Affymetrix) and γ-^{32}P-ATP (Perkin-Elmer) following manufacturer's instructions, and ethanol precipitated.

Construct 1: 5' – GGT TGG CGC GAA GCA TTC GCG GGT TGG - 27 nt
Construct 3: 5'- GGG AAG GGC GCG AAG CAT TCG CGA GGT AGG -30 nt
Construct 5: 5'- TTG GGT GGG CGC GAA GCA TTC GCG GGG TGG GT- 32 nt
CatG4: 5' - TGG GTA GGG CGG GTT GGG AAA - 21 nt

3.3.2. UV-Vis spectroscopy

DNA stock solutions were diluted in reaction buffer (final concentration of 40 mM HEPES pH 8.0, 20 mM KCl, 1% dimethyl formamide, 0.05% Triton X-100) and mixed with 1 μM hemin. Samples were incubated for 30 minutes before being transferred to 500 μL cuvettes for measurement on a ThermoFisher Evo 260 UV/Vis spectrophotometer. Hemin was kept constant at 1 μM and titrated with DNA. K_d was calculated by fitting the equation Y = Bmax*X/(Kd + X)

3.3.3. G & T ladders

G ladders were created by mixing radiolabeled DNA (~2000 cps) into a 20 μL reaction containing 100 mM sodium cacodylate and 1 μM unradiolabelled DNA. For

samples containing salt, KCl (20 mM) and $MgCl_2$ (5 mM) were added and allowed to fold for 5 minutes. Addition of 1 µL of 2% DMS initiated the methylation reaction and was left for 30 minutes at room temperature. DMS methylates the N7 position of guanine bases in single and double stranded DNA, which are then susceptible to base cleavage. The reaction was quenched with 2 µL of 2-β-mercaptoethanol, and then ethanol precipitated. Samples were resuspended in 100 µL 10% piperidine solution, to cleave the DNA backbone at sites of modification, and placed at 90°C for 30 minutes before being vacuum dried. Samples were mixed with denaturing loading dye prior to running on a gel.

T ladders were created by adding 4 µl of 160 mM $KMnO_4$ solution to radiolabeled DNA (2000 cps) in a total volume of 40 µl (containing no salt, or 20 mM KCl, or 20 mM KCl and 5 mM $MgCl_2$). $KMnO_4$ selectively oxidizes single stranded thymine bases, which then become susceptible to base cleavage. Samples were incubated at RT for 2 minutes before quenching with 2 µL 2-beta-mercaptoethanol, followed by ethanol precipitation. Samples were resuspended in 100 µL 10% piperidine solution and placed at 90°C for 30 minutes before being vacuum dried. Samples were mixed with denaturing loading dye prior to running on a gel.

3.3.4. Gel electrophoresis

Samples were loaded into 12% denaturing polyacrylamide gels. After gels were run for 20 minutes to 3 hours, they were placed in plastic wrap and placed on phosphorous screens (Amersham Biosciences) for various exposure times (1–24 h), at -20ºC. The visualization of radioactive gel bands was done by scanning screens in a Molecular Dynamics Typhoon 9410 Variable Mode Imager. The gel images were analyzed and quantified using ImageQuant 5.2 (GE Healthcare).

3.3.5. Hemin footprinting ladder

DNA stock solutions were diluted to 1 µM in reaction buffer (final concentration of 40 mM HEPES pH 8.0, 20 mM KCl, 1% dimethyl formamide, 0.05% Triton X-100), and heated to 95°C for 5 minutes, then cooled to room temperature. Hemin was added to a final concentration of 5 µM, allowed to bind and then the reaction was initiated by addition of H_2O_2 at 500 µM. The reaction was allowed to proceed for 30 minutes before stoppage

by ethanol precipitation, followed by 10% piperidine treatment, vacuum drying and resuspension in loading dye prior to gel running.

3.4. Results and Discussion

3.4.1. Hemin binding characteristics of duplex quadruplex hybrid structures

To determine if the selected duplex-quadruplex hybrid structures bind hemin, a titration of hemin with DNA was done. Hemin-quadruplex interactions show a characteristic hyperchromicity at the 398 nm peak of the hemin absorption spectrum.[39] All of the selected hybrids bind hemin, demonstrated in Figure 3-2 by the characteristic red shift and hyperchromicity of the Soret peak at 398 nm in the hemin-containing spectra. Panel A, representing a CatG4 (parallel quadruplex derived from PS2.M) titration of hemin, shows the expected binding of hemin to a quadruplex.[39] In a parallel quadruplex, it is common to see the Soret peak shift from ~398 to 404-405 nm. The calculated K_d of CatG4 was found to be 0.7 μM, as seen in Figure 3-3. Construct 1-hemin absorption spectra show a similar hyperchromicity to CatG4, but differ in the extent of its red shifting (λ_{max}=407 nm). The K_d of the antiparallel Construct 1 (Con1) was determined to be 1.7 μM. The mixed orientation Construct 3 (Con3) was a weaker binder and did not have a clear K_d for the range tested. Construct 5 (Con5), adopting a fully parallel structure, shows a K_d of 1.4 μM, and a similar binding pattern to the parallel CatG4. In the cases of Con1 and Con5, the additional duplex has relatively little effect on the ability of hemin to bind.

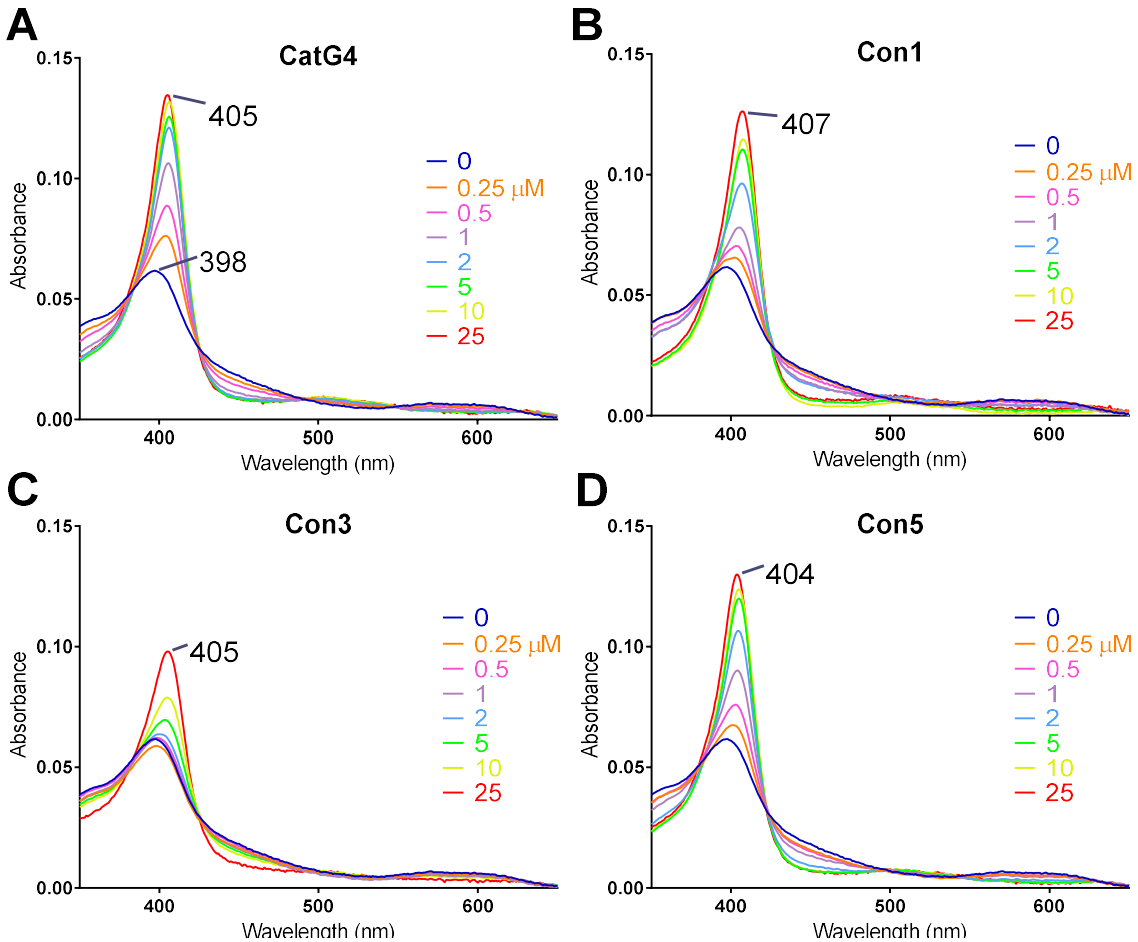


Figure 3-2 **Hemin absorption spectra upon quadruplex interaction. Hyperchromicity at the Soret (~400) is indicative of binding. Traces are averaged spectra of 2 samples except for: all 25 µM samples, Construct 1 - 2 and 10 µM samples.**

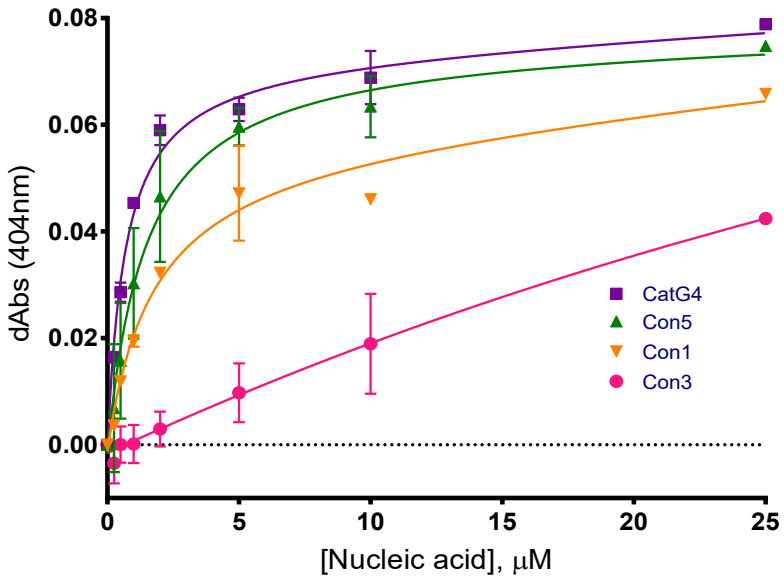


Figure 3-3 **Change in absorbance of hemin at 404 nm as a function of DNA concentration. Fitting of curves yields dissociation constants (K_d). Error bars represent s.d. of 2 repetitions. Lack of error bars represents 1 repetition. Equation used does not account for ligand depletion. Values are indicative of binding in relation to CatG4.**

3.4.2. Salt effects and hemin footprinting

Guanine and thymine ladders were created for Construct 1, 3 and 5 to observe protection patterns that would identify folding in different regions. Creating protection patterns or damage patterns of DNA bases with chemical treatments and visualizing them on polyacrylamide gels is often referred to as a ladder or footprint. These reactions can be chosen to be specific to certain bases and chemical modification is dependant on the solvent accessibility of those bases. Changing the salt conditions that the footprinting reactions are conducted in changes what structures are supported. This has the affect of protecting certain bases more or less from chemical modification. By treating the chemically modified DNA with base and heat, sites of modification become locations of backbone cleavage, which can be observed when we run the samples in a polyacrylamide gel. By comparing these footprints with the NMR structures of the constructs determined by Phan and coworkers[62] we can make inferences about structural changes. The NMR structures were obtained in only 20 mM K^+ [62]. The addition of Mg^{2+} is used to see whether stabilizing the duplex region would affect the quadruplex. Mg^{2+} stabilizes duplex DNA by

complexing the phosphate backbone and reducing the high localization of negative charge.

Figure 3-4 shows one of the NMR structures for Construct 1 and a simplified representation of Construct 1, with the G and T ladders below. Guanine footprinting by DMS selectively methylates guanine bases at the N7 position, making them labile when treated with base. Protection from DMS modification indicates protection at this N7 position, as is present in the Hoogsteen bonding patterns of quadruplex structures. In the no salt condition, all guanines show a relatively equal level of damage. The addition of either K^+ or Mg^{2+} to the guanine footprinting reaction had minor effect on the overall guanine damage pattern. Quadruplex guanines show less damage and greater protection upon the addition of K^+ salt, which is consistent with the formation of more stable secondary structure. Guanine 21 of Construct 1, sitting at the crossover between duplex and quadruplex, is consistently unprotected and shows increased damage upon the addition of K^+ and Mg^{2+}, unlike other duplex guanines that are more consistent in their patterns of damage. As K^+ and Mg^{2+} stabilize the quadruplex and duplex regions respectively, G21 becomes less protected, likely due to increased strain.

Thymine footprinting creates oxidative damage selective mostly for single stranded thymines by treating the sequence with $KMnO_4$. In construct 1, only T24 and T3 show significant damage which increases with the addition of salts. This indicates that they are the only thymines not engaging in significant bonding, consistent with them being in the loop regions. The NMR data suggest T25 and T4 are involved in hydrogen bonding and may explain why they are protected[62]. The duplex thymines are protected, indicating a stable duplex.

Hemin footprinting has been useful in the past to determine where on the quadruplex the hemin is bound.[43] When a G-quadruplex-hemin complex is reacted with H_2O_2 in the absence of an oxidizable substrate, oxidation of guanine bases begins to occur. Oxidation of quadruplex guanines commonly occurs at positions nearer to the site of hemin binding. In parallel quadruplexes like CatG4, where both outward facing G-quartets present a similar binding surface, the pattern of oxidative damage is symmetrical and it occurs at all 4 positions of the G-quartets. Figure 3-4 demonstrates the application of hemin footprinting on Construct 1 (bottom right gel image). In Construct 1 there are 2 major locations of damage: G21 and G8. These two duplex guanine positions suggest that

either the majority of binding and hemin activation is occurring at the one side of the quadruplex, or that these bases have a higher oxidation potential and electron transfer from other positions occurs prior to the formation of irreversible oxo-guanine species. It may also be related the solvent exposure of these duplex positions. Construct 1 is a mixed orientation quadruplex, which makes one binding surface crowded with loop thymines and may explain the preference. The duplex-quadruplex junction also crosses and crowds the opposite G-quartet surface in Construct 1, but the NMR structure suggests the major groove of the duplex is oriented in such a way that G1, G6, and G27 of the quadruplex remain accessible to binding hemin. The duplex guanines G21 and G8 are positioned over the G-quartet, theoretically putting them near to a hemin's axial coordination site. The specific damage pattern of Construct 1 presents an interesting case of a quadruplex with a more defined binding "pocket." Oxidative damage appears to be consistently targeted to 2 duplex guanines, as opposed to all quadruplex guanines. It may suggest an opportunity for enantio or regioselective oxidation of substrates that is unavailable in regular quadruplex-hemin complexes.

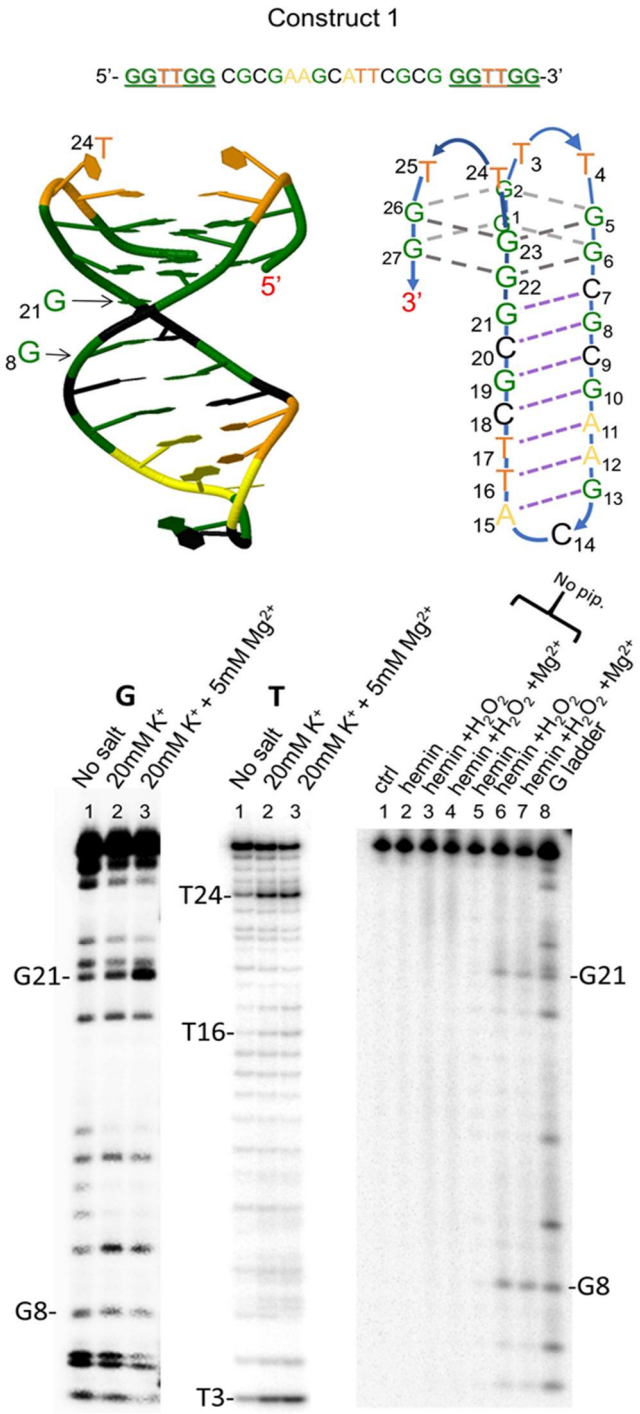


Figure 3-4 **Construct 1 NMR structure (PDB:2M8Z[62]) next to cartoon representation (top). Bases in both are color coded to match. G, T and hemin footprinting ladders (bottom). Treating the sequences with DMS (G ladder) and $KMnO_4$ (T ladder) in combination with different salt conditions can show participation of individual bases in secondary structures. Hemin footprint identifies guanines that become oxidized in a peroxidase reaction lacking substrate, suggesting proximity to the hemin binding site. Gel patterns and trends were supported by 2 repetitions.**

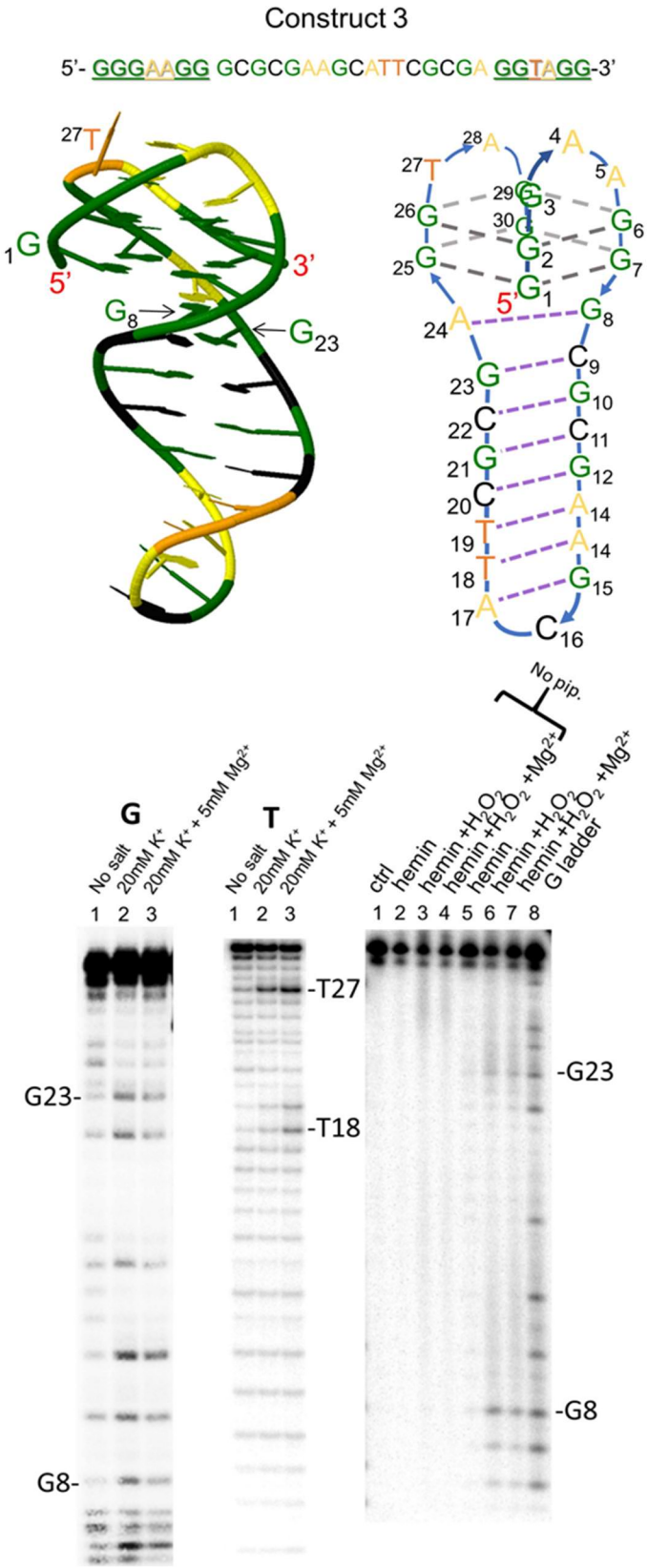


Figure 3-5 **Construct 3 NMR structure (PDB:2M91[62]) next to cartoon representation (top). G, T and hemin footprinting ladders (bottom).**

The G ladder for Construct 3 (Figure 3-5), an antiparallel quadruplex, shows protection of the 3' G-rich regions upon addition of K^+ and Mg^{2+}, but less so for the 5' region. The reduced protection of the 5' guanines (G1, G2, G3, G6, G7) indicates less stable incorporation of these bases into the quadruplex. The G-quadruplex could form using G1 and G2 in the quartets or G2 and G3. The T ladder damage pattern shows increased vulnerability upon the addition of salts, specifically at T27 and T18-19. T27 is a loop thymine and stabilization of the G-quadruplex should lead to increased damage. T18-19 are duplex positions and the addition of Mg^{2+} has the effect of increasing the vulnerability to damage. This indicates that duplex integrity near the stem loop region is sacrificed for stability overall.

The hemin footprint of Construct 3 indicates damage is concentrated on the duplex facing surface of the quadruplex, as in Construct 1. Again, the loops may obscure the opposite binding face. The two duplex guanines nearest the G-quadruplex, G23 and G8, receive the majority of damage, with G6 and G7 showing minor damage. The addition of Mg^{2+} has little impact on the pattern of damage and indicates that the state of the duplex near to the loop has no correlation to hemin activation.

The G ladder for Construct 5 (Figure 3-6) shows well protected quadruplex guanines upon addition of K^+ and no change of protection upon addition of Mg^{2+}. G30 and G31 show some signs of susceptibility to damage. In the T ladder, T28 and T6 increase in damage in the presence of K^+ and Mg^{2+}, as they are unpaired in loops of the quadruplex. Duplex T19 and T20 increase in damage significantly in the presence of Mg^{2+}, as observed for Construct 3. This observation points to a destabilization of base pairing near the stem loop while the rest of the duplex is stabilized.

The hemin footprinting in Construct 5 is indicative of multiple binding sites, as would be seen in a non-hybrid parallel G-quadruplex. There is greater damage at G31 and G27 than at their counterparts on the opposite face of the quadruplex, G29 and G25. This indicates a hemin binding preference or greater activation of hemin on the 3' side, or possibly a lack of nearby duplex guanines to intercept the damage. Damage at nearby non-quadruplex guanines G24 and G11 is observed, although the duplex is no longer positioned as directly above the binding site. The presence of unpaired bases at the duplex-quadruplex junction may allow for flexibility in the duplex allowing these guanines

to move closer to the catalytic zone. In the NMR structure, an unpaired G24 can be seen to be positioned above the 5' face.

3.5. Conclusion

Duplex-quadruplex hybrids represent a stable secondary structure of DNA that have analogous sequences in the human genome. The data shown here identify their ability to bind hemin and activate peroxidase activity. Constructs 1 and 5 show binding affinities similar to that of parallel quadruplex CatG4 based off of changes in the absorbance spectra of hemin. The lower hyperchromicity observed in Construct 3 when it is mixed with hemin may be a result of lower binding affinity. It is also possible the attached duplex is influencing the absorbance change brought on by the binding interaction. The addition of Mg^{2+} has only minor effects on the stability of the quadruplex and duplex regions versus K^{+}. Hemin footprinting in Constructs 1 and 3 suggests that hemin binding is not prevented by the attachment of a duplex across or adjacent to a G-quartet. Further experiments are needed to confirm these binding locations, as well as to determine if the duplex portions affect the oxidations of substrates. Antibodies targeting G-quadruplexes have been effective for *in vivo* visualization experiments[38]. Raising an antibody against a duplex-quadruplex hybrid structure may be useful to determine if such structures are viable *in vivo*, however the variability in loop length of these structures may make this difficult. The research presented in Chapter 4 may have relevance for identifying non-canonical quadruplex morphologies that are capable of binding and activating hemin.

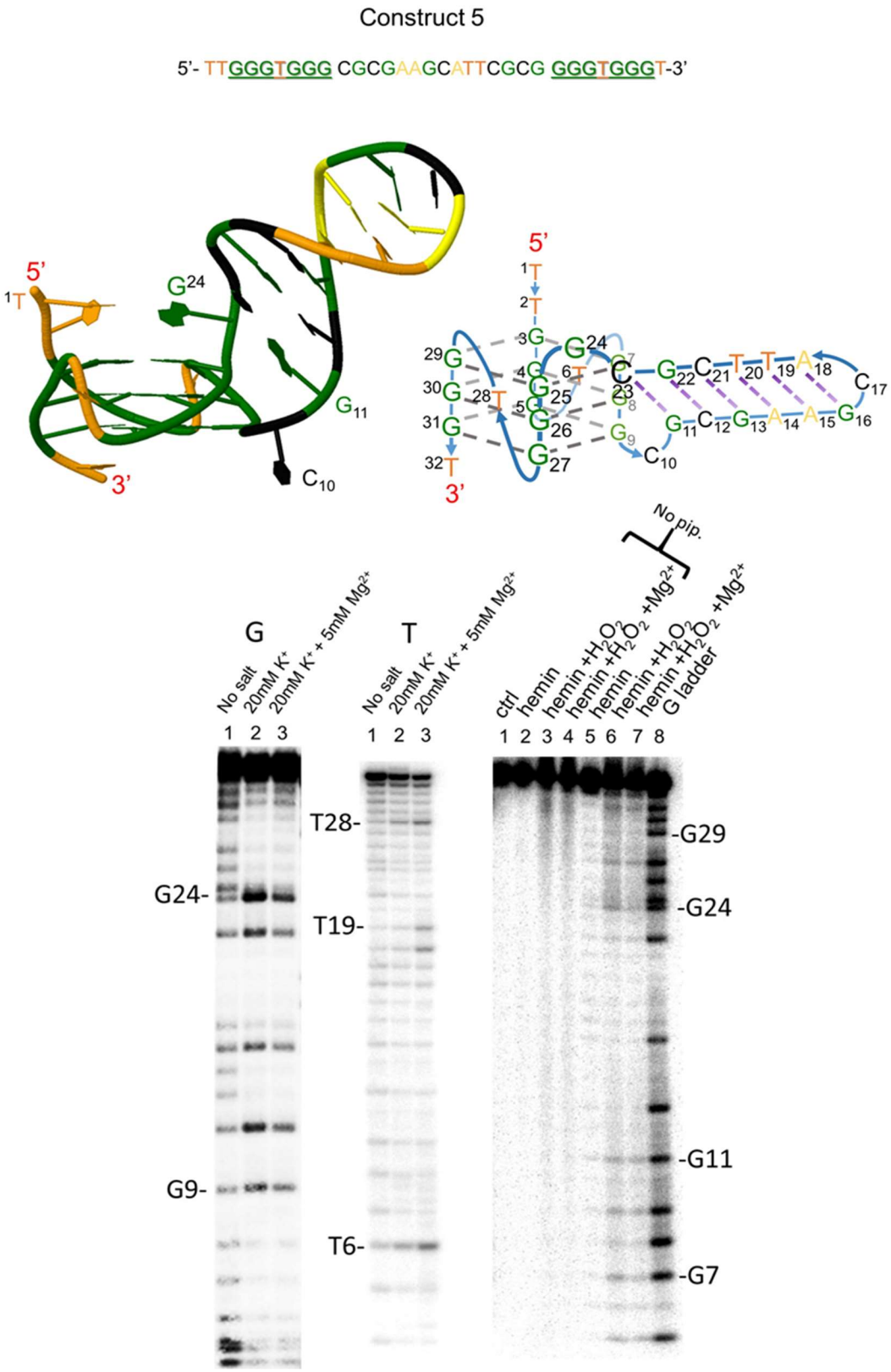


Figure 3-6 **Construct 5 NMR structure (PDB:2M93[62]) next to cartoon representation (top). G, T and hemin footprinting ladders (bottom).**

Chapter 4.

G-Quadruplex tagging through intrinsic peroxidase activated biotinylation

4.1. Introduction

The structural and functional relationships of G-quadruplexes within the cell have been difficult to characterize. The recent successful visualization of G-quadruplexes in live cells through the use of fluorescent antibodies highlights some temporal and spatial patterns.[33] To further understand the mechanisms involved, and to identify genomic sequences involved in forming G-quadruplexes, techniques such as ChIP-Seq have been employed to map natural quadruplex protein binding locations.[66] The ability of ChIP-Seq to crosslink proteins and DNA in physical contact makes it a valuable for targeting quadruplex interactions. For example, Pif1 DNA helicase, known to unwind quadruplexes *in vitro*, was used in a ChIP-Seq on the *S. cerevisiae* genome to successfully target predicted quadruplex motifs.[67] Another group used a multi-step approach, that included the known small molecule G-quadruplex binder and stabilizer, pyridostatin (PDS), to form and/or stabilize G-quadruplexes within living cells.[68] A PDS-mediated G-quadruplex formation results in replicative pausing and triggers a DNA damage response. One of the proteins triggered by the response was then used as a target for ChIP-Seq analysis. In this study, regions of genomic DNA with higher algorithmically predicted G-quadruplex formation showed increased representation in the analysis.

Higher resolution methods to map quadruplexes have more recently been described.[69] A 2015 study combined Illumina high throughput sequencing of the human genome with a polymerase stop assay under conditions favouring quadruplex formation. Potential quadruplex forming sites were categorized based upon data from 2 reads. The first read proceeds normally in Na^+ and then the template and primer sequence are melted such that the template can be reprimed. A second read is done under quadruplex supporting solution/ionic conditions, and if the decrease in sequencing quality is deemed to be sufficiently large those regions are annotated as “potentially quadruplex forming.” The quadruplex-supporting conditions used in this study were the presence of 50 mM K^+ or of 1 μM PDS. PDS treatment yielded 716,310 sites, and K^+ showed 525,890 sites in the

genome. These present a significant increase over the numbers predicted by the previously mentioned predictive algorithms, which identified ~361,000 potential quadruplexes in the human genome.[64] The sequences identified by the PDS-treatment condition incorporated 73% of the predicted 361,000 quadruplex sequences identified by the predictive algorithm, whereas those from the K^+-treatment condition incorporated 60% of the sequences highlighted by the predictive algorithm.

The abundance of potential quadruplex forming sequences in the human genome raises many questions with regard to function. There is growing evidence to suggest that G-quadruplexes may play roles in certain diseases. Work in our lab, and elsewhere, has started to investigate neurodegenerative conditions featuring abnormal DNA repeat expansions, such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). An expansion of the sequence GGGGCC in the *C9orf72* gene has been found in various forms of these diseases. Repeats of this sequence can form G-quadruplexes that are capable of activating heme *in vitro.*[56] A common factor in many neurodegenerative diseases is the presence of oxidative stress. Developing a new way to locate G-quadruplexes that form *in vivo* and have the potential to be catalytically activated would be of great value. The research presented in this chapter will focus on the creation of an assay towards this goal.

4.2. Phenolic peroxidase substrates

In situ hybridization (ISH) and immunohistochemistry (IHC) protocols allow for visualization of targets through the use of labelled nucleic acid probes and fluorescently labeled antibodies. The specificity of these probes allows for high resolution imaging of the target molecules, but high background staining presents a key challenge. One remedy for this is the tyramide signal amplification assay, which is used to amplify the signal provided from ISH and IHC.[70] In this procedure, the primary antibody or probe does not generate an output signal in itself. Instead, imaging is achieved through the addition of a secondary antibody that is conjugated to the enzyme horseradish peroxidase (HRP); reactive HRP substrates serve as probes. When such reactive, often phenolic, substrates (such as derivatized tyramides) are added in the presence of H_2O_2, those in the proximity to the HRP antibody react to form short lived (lifetime <1 ms) free radical tyramide intermediates.[71] Such highly reactive intermediates quickly react covalently with nearby proteins and nucleic acids. If the tyramide probe is fluorescent, regular visualization

methods are then enabled, with 10-100X signal amplification with little loss in resolution. If the tyramide probe was biotin derivative, streptavidin pulldowns may be performed to identify the molecules present at the targeted location.

This basic premise has been adapted elsewhere with success. Ting and coworkers engineered an ascorbate peroxidase (APEX2) that could be targeted to a specific region of a living cell, and upon addition of H_2O_2 and a membrane permeable biotin-containing tyramide probe, they could covalently biotin-tag proximal proteins to be pulled down or visualized.[72] With a fluorescent tyramide probe, they were able to visualize proteins specific to the inner mitochondrial membrane within a live cell. By using a biotinylated probe, they achieved purification and characterization of the proteome specific to the mitochondrial matrix as assayed using mass-spectrometric analyses of pulled down cellular proteins. They were able to obtain unique western blotting patterns of biotin tagged proteins when APEX2 was conjugated to nuclear, cytosolic, mitochondrial matrix, and intermembrane space proteins.[73] Furthermore, conjugating APEX2 with different targeting sequences for the same compartment led to similar western blotting patterns. This techniques compatibility with fluorescence, staining, and mass spectrometric methods allows for the ability to analyze protein localization and identity on a variety of scales in living cells.

In the above protocol the peroxidase APEX2 is targeted to a region of interest through genetic fusion to targeting peptide sequence, or a protein known to be specific to the desired cell compartment[73]. The cells are incubated 30 minutes in cell medium containing 500 μM of the biotinylated tyramide compound. To initiate the reaction, 1 mM H_2O_2 is added and quenched after 1 minute. Cells are fixed with paraformaldehyde prior to imaging or lysed and gel purified for further enrichment with streptavidin beads.

Important factors to consider when dealing with this tagging method are the radius of labeling and the location of labelling. The 1 minute window of labeling allows for proteins tagged with biotin to diffuse away from the original site if it is not membrane enclosed, potentially leading to spurious signals. The radius of labelling for the tyramide radical itself has been shown to be as small as 20 nm in electron microscopy experiments on fixed cells, but would likely vary depending on the local concentration of radical scavengers.[73] Ting and coworkers also analyzed the membrane permeability of a tyramide radical first by targeting APEX2 to the inner wall of the cellular membrane and then adding an alkyne

tyramide and H_2O_2. By then running a dipolar cycloaddition “click” chemistry assay with a membrane impermeant fluorescent azide, they identified that exterior cell surface proteins were not labeled with alkyne due to the lack of fluorescence signal. The same result was found when biotin-tyramide was used with impermeant streptavidin. The phenolic tyramide radical is known to favour reactions with tyrosine residues, coupling with the phenolic side chain.[74] It is also likely that other electron rich amino acids like tryptophan, cysteine and histidine are tagged.[73] There is therefore some inherent selectivity when labeling proteins with these accessible residues.

We hypothesized that a similar tyramide based approach could be used to label DNA as there is evidence that suggests a phenoxyl radical will attack the C8 position on guanine bases.[75] Ochratoxin A (OTA) is a mycotoxin containing a chlorophenolic moiety that is known to induce oxidative stress and DNA damage in cells. The reactive intermediate involved is thought to be the OTA phenoxyl radical.[76] Oxidation of OTA with horseradish peroxidase and H_2O_2 in the presence of deoxyguanosine led to the isolation of OTA-dG adducts. NMR and MS analyses identified the position of the bond from the phenol group of OTA to the C8 position on dG.

The peroxidase activity of the G-quadruplex-hemin complex offers a unique opportunity to explore a tyramide labeling reaction, with the goal of tagging peroxidase active quadruplexes or nearby proteins for a proteomic analysis. We obtained a biotinylated tyramide (Figure 4-1) with the intention of investigating the validity of a G-quadruplex proximity tagging assay.

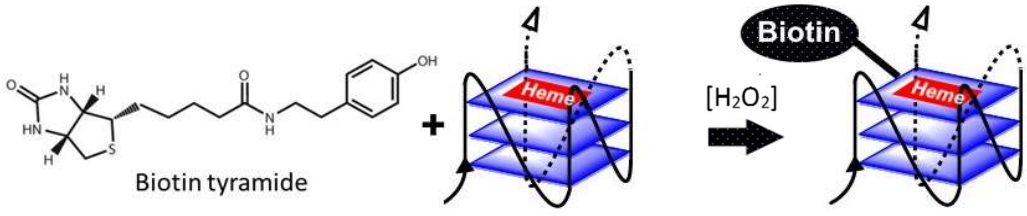


Figure 4-1 Scheme of biotin tyramide and hemin-quadruplex reaction. Tagging is shown on quadruplex but may occur elsewhere.

4.3. Materials and Methods

4.3.1. Oligonucleotide purification

All oligonucleotides were purchased from Integrated DNA Technologies (IDT). The samples were solubilized in a 10% piperidine and placed at 90°C for 30 minutes to cleave damaged oligos. After vacuum drying, the oligos were gel purified in 8-12% denaturing gels followed by extraction by crushing the gel into small pieces and soaking in TE buffer, followed by butanol extraction, ethanol precipitation and resuspension in TE buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0). For gel visualization, DNA was ^{32}P-labeled using OptiKinase (Affymetrix) and γ-^{32}P-ATP (Perkin-Elmer) following manufacturer's instructions, and ethanol precipitated.

The following sequences were used in the biotin tyramide reaction.
Construct 1 (from Ch. 3): 5' – GGT TGG CGC GAA GCA TTC GCG GGT TGG - 27 nt
CatG4: 5' - TGG GTA GGG CGG GTT GGG AAA - 21 nt
duplex (w/ reverse complement): 5' – TTT AGC TCA CGA GAC GCT CCC ATA GTG A - 28 nt

The CatG4 parallel quadruplex was extended at both ends with non-specific bases to create this sequence for use in the Fenton reaction.
Fenton G4: 5' - ACA TAG CTG ACT GGC TTG ATT TTG GGT AGG GCG GGT TGG GAA ATA TCG AAT TCT CAG CCT ACA CTG CAG TAC TA - 74 nt

The following sequences were used in the PCR experiments. The PCR G4 template is an extended version of CatG4 with primer sequence at each end. The reverse primer binding site is nested so the PCR product is shorter than the template.
PCR G4 template: 5' - ACA TAG CTG ACT GGC TTG ATT TTG GGT AGG GCG GGT TGG GAA ATA TCG AAT TCT CAG CCT ACA CTG CAG TAC TAG TAC ATA TCA – 84 nt
PCR reverse primer: 5' – TAG TAC TGC AGT GTA GGC TG - 20 nt
PCR forward primer: 5' – ACA TAG CTG ACT GGC TTG AT - 20 nt

4.3.2. Biotin tyramide reaction

Biotin tyramide (Toronto Research Chemicals) was solubilized in dimethyl sulfoxide as a 100 mM stock. Hemin (Frontier Scientific) stock solutions were prepared fresh in dimethyl formamide to 1 mM. H_2O_2 solutions were diluted from a 50% stock to 100mM in ddH_2O. DNA stock solutions were diluted in reaction buffer (40 mM HEPES pH 8.0, 20 mM potassium chloride, 1% dimethyl formamide, 0.05% Triton X-100), heated to 95°C for 5 minutes and then cooled to room temperature. Hemin and biotin tyramide were added, followed by the initiation of the peroxidase reaction by addition of H_2O_2. Reactions were allowed to proceed for 30 minutes, followed by the quenching of the reaction by ethanol precipitation or addition of catalase enzyme (if indicated). For polyacrylamide gel loading the samples were resuspended in streptavidin solution (20 mM KH_2PO_4 buffer) for 5 minutes, followed by the addition of gel loading buffer (denaturing buffer:10mM EDTA in 95% formamide, xylene cyanol, bromophenol blue; native buffer: 50% glycerol, 2.5X TBE, bromophenol blue). For footprinting analyses, biotin tyramide treated samples were resuspended in 10% piperidine and heated at 90°C for 30 minutes, vacuum dried and then streptavidin and loading dye were added. After gel running, polyacrylamide gels were placed on phosphorous screens (Amersham Biosciences) for various exposure times (1–24 h), at -20ºC. The visualization of gel bands was done by scanning screens in a Molecular Dynamics Typhoon 9410 Variable Mode Imager at a 200 micron pixel size. Some digitalization error is incurred through this process. The gel images were analyzed and quantified using ImageQuant 5.2 software (GE Healthcare).

4.3.3. Fenton reaction

A solution of 20 mM iron (II) stock was prepared by dissolving $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ in ddH_2O. This is mixed with a 2-fold excess of EDTA to ensure all iron (II) is bound to EDTA. A 10 mM stock of sodium ascorbate is prepared. H_2O_2 is prepared at 0.6% wt/vol. A stock of stop buffer is prepared as a 100 mM thiourea. After samples of DNA had undergone 5' radiolabeling with ^{32}P and then been exposed to the biotin tyramide assay, they were exposed to the Fenton conditions. A 1 µL sample of DNA is mixed with 1 µL of each of the reagents: 1 mM Fe(II)/2 mM EDTA solution, 10 mM ascorbate, and 0.6% H_2O_2. Reactions were quenched at various time points by addition of 1 µL of the thiourea stop buffer. Samples were then ethanol precipitated and sequenced

on a 12% denaturing gel. The reaction yields non-specific cleavage of the DNA sequence via hydroxyl radicals generated through the reaction $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + \bullet OH + OH^-$

4.3.4. PCR

DNA samples treated or untreated with the biotin tyramide assay were streptavidin shifted and gel purified to form various categories of shifted or unshifted bands. Samples from each of these categories were then subjected to standard a standard PCR protocol. Each sample contained ~20 pmol of DNA was added to a 20 µL PCR reaction containing 1X Taq reaction buffer, 1 U of Taq polymerase enzyme, and 0.2 µM forward and reverse primers, and 200 µM dNTP's. Control samples were made without the forward primer or were at room temperature during the cycling. Cycling conditions were: initial denaturation- 30 secs @ 95°C; 20 cycles of 95°C for 30 secs, 52°C for 30 secs; 68°C for 45 secs; final extension- 2 mins @ 68°C. The samples were then ethanol precipitated and run on a 10% denaturing gel.

4.4. Results and Discussion

4.4.1. Biotin tyramide reaction

It was first necessary to determine if a hemin-quadruplex DNAzyme could oxidize the phenolic substrate biotin tyramide. To do so, hemin, ^{32}P labeled quadruplex, H_2O_2, and biotin tyramide were mixed, followed by the addition of the biotin-binding protein streptavidin. The assumption was made that if biotin tyramide was oxidized, some of the phenoxy radicals should covalently react with the quadruplex. The portion of quadruplexes that were labeled could then be distinguished from unlabeled quadruplexes through the addition of streptavidin, which would bind the biotin labeled quadruplex. This could be visualized by observing a change in gel mobility, as any streptavidin bound fraction should be much larger and run more slowly through the gel. The biotin-streptavidin interaction has been shown to be stable even under the highly denaturing conditions of the loading buffer and the gel itself.[77]

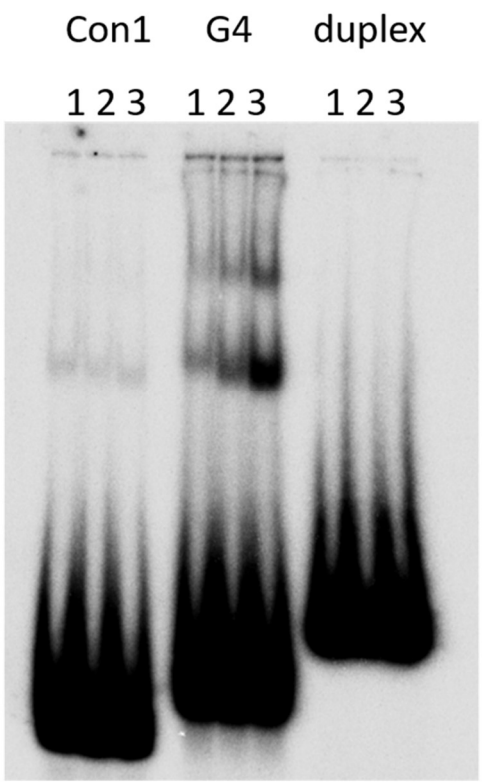


Figure 4-2 **Biotin tyramide reaction on Construct 1, CatG4 (G4) and a duplex control. 10% denaturing gel. Lanes labelled 1 contain 5 μM hemin, 1 μM DNA, 1 mM H_2O_2 and 60 μM streptavidin. Lanes 2 and 3 contain same as 1 plus 5 and 50 μM biotin tyramide respectively. The smearing is likely salt effects, as a result of a direct loading of the non-ethanol precipitated assay into the gel. Subsequent gel analysis used precipitated samples to avoid this complication.**

The results from the initial reaction of quadruplex-hemin with H_2O_2 and biotin tyramide are presented in Figure 4-2. The first set of reactions included antiparallel G-quadruplex Con1, parallel CatG4 (simplified to G4), and a duplex control. Reactions loaded in lanes labelled 1 in Figure 4-2 were intended to serve as a negative control, as they lacked biotin tyramide and we would expect no streptavidin shifting. Lanes labeled 2 and 3 contain the full mix of reactants, with the addition of biotin tyramide at 5 μM and 50 μM, respectively. Five μL of the streptavidin stock were added to 5 μL of the reaction mixture after 30 minutes, with no quenching reagent. This lack of quenching turned out to have implications for background reactions. To this, an equal volume of denaturing loading dye was added (10mM EDTA in 95% formamide), and 5 μL of this was loaded into each lane of the gel.

The duplex shows no signs of tyramide activation, with no lanes showing a streptavidin shift. Construct 1 shows signs of streptavidin shifting in all 3 lanes that remain constant upon addition of biotin tyramide. The presence of streptavidin shifting in the absence of tyramide suggests another mechanism for the tagging of quadruplex with streptavidin. The CatG4 sample shows biotin tyramide-dependent streptavidin shifting pattern, with the 50 μM biotin tyramide treatment showing a larger percentage shift of radiolabeled DNA than 5 μM. The 0 μM tyramide condition also shows shifting, and to a greater extent than that in the Construct 1 case, suggesting that the same background tagging mechanism is present in both, but to a greater extent in the parallel G-quadruplex, CatG4. As previously described, the hemin-G4 peroxidase reaction is known to produce a carbon centered radical, likely based on the quadruplex.[43] The lack of quenching before addition of streptavidin could have allowed these radicals to directly react with streptavidin. Con1's inability to activate the substrate tyramide, while still activating the background reaction may be an indication that the duplex region of the duplex-quadruplex hybrid is interfering with biotin tyramide oxidation. This could be a result of the size or shape of the substrate, or an indication that the reactivity of the hemin-quadruplex is preferentially directed toward oxidation of the duplex guanines mentioned in Chapter 3.

The observation that Con1 was not appreciably labeled in a predictable manner led to a focus on CatG4 for the remainder of the biotin tyramide assay experiments. Figure 4-3 shows an 8% denaturing gel of a preparative tyramide assay with CatG4. The more resolved bands in this gel can be attributed to the ethanol precipitation of samples before addition of streptavidin and loading dye. The obvious streptavidin shifting pattern is apparent in lane 5. Two distinct low mobility bands suggest either 1 streptavidin molecule bound to 1 or more tagged DNA molecules, or multiple biotinylations on a single CatG4 molecule leading to the binding more than 1 streptavidin molecule. This could be clarified with mass spectrometry. The size separation of the two shifted bands supports multiple biotinylations on a single CatG4, as later gels with nucleotide ladders show this distance to be greater than the 21 nucleotides present in CatG4. The smearing pattern in lane 4 also indicates the likely due to the production of biotinylated forms of CatG4 running with different mobility. An absence of this smear in lane 5 suggests most forms are capable of binding streptavidin. Addition of streptavidin to the biotin tyramide treated sample also leads to an increased proportion of CatG4 that remains in the well at the top of the gel.

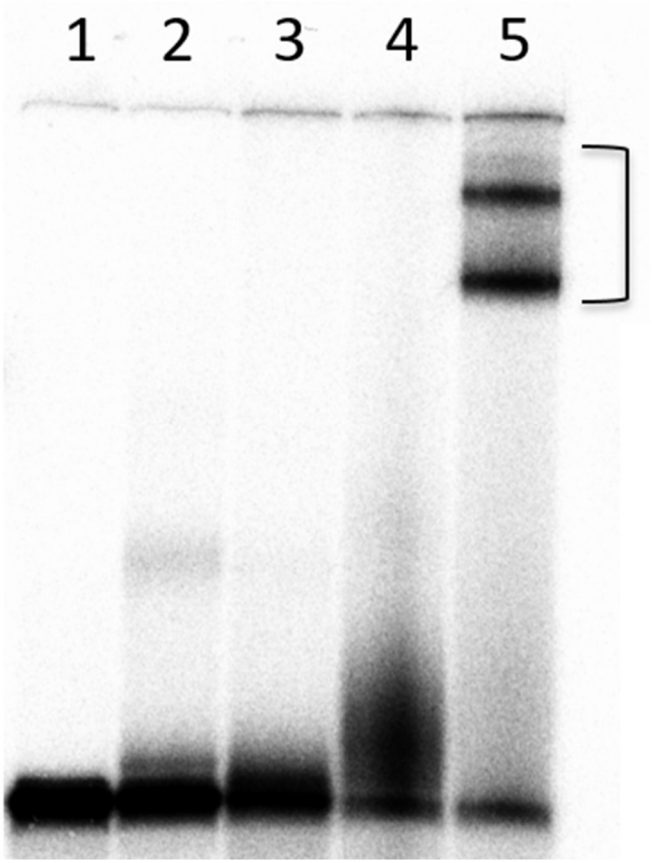
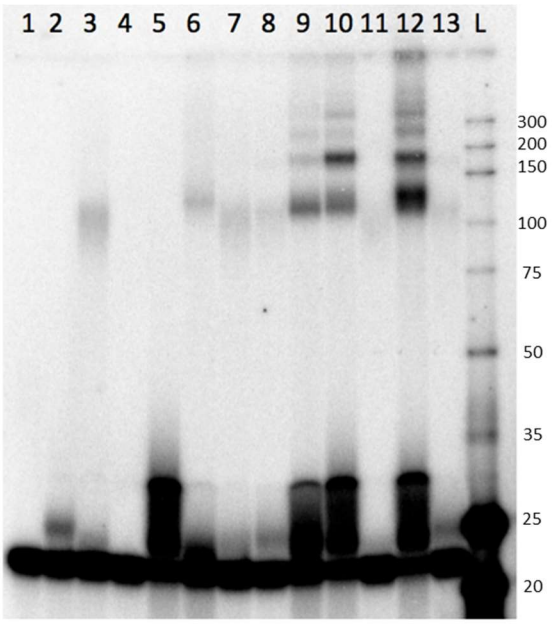


Figure 4-3 **Biotin tyramide assay on CatG4. Additional purification and quenching steps were added to reduce smearing. 10% denaturing gel. Lane 1 contains CatG4 alone; lane 2 contains CatG4 and hemin; lane 3 contains CatG4, hemin and H_2O_2; lane 4 contains CatG4, hemin, H_2O_2 and biotin tyramide; lane 6 contains CatG4, hemin, H_2O_2, biotin tyramide and streptavidin. 10 µM CatG4, 50 µM Hemin, 200 µM biotin tyramide, and excess streptavidin.**



	1	2	3	4	5	6	7	8	9	10	11	12	13
G4 DNA	+	+	+	+	+	+	+	+	+	+	+	+	+
hemin		+	+		+	+	+	+	+	+	+	+	+
StAv	+		+	+		+	+	+	+	+	+	+	+
Bio-Tyr (μM)		50		50	50		5	50	5	50		50	50
H_2O_2					+	+			+	+		+	
Salmon DNA	+	+	+	+	+	+	+	+	+	+			

Figure 4-4 Biotin tyramide reaction with controls. 12% denaturing gel. Reactions contained 1 μM hemin, and 1 μM DNA, 500 μM H_2O_2, 100X weight excess of salmon sperm DNA and the listed amount of biotin tyramide.

Figure 4-4 represents a more systematic analysis on the effect of varying the reaction conditions. The 20 μL reactions had 5 μL of streptavidin added (~50 μM), and were left for 5 minutes. One μL of this sample was then mixed with 5 of loading dye before heating 2 minutes at 90° C and loading. In order to minimize smearing a smaller amount of the reaction was loaded with a greater proportion of loading dye. The higher resolution of this 12% gel identifies that the streptavidin banding pattern does contains more than 2 bands, as seen most prominently in lane 12. In this lane, there are at least 4 distinct bands induced by the addition of streptavidin (located above 100 nt on the DNA ladder), and 2 more faint bands appear to be present above that. The addition of salmon sperm DNA

(~600 bp fragments) does have an effect on the amount of biotin labelled G-quadruplex, seen in lanes 10 versus 12, where the percentage of streptavidin shifted bands make up 13% and 27% respectively.

Lanes 1 and 4 show a lack of any shifting behavior in the absence of hemin. Lanes 2 and 3 indicate that having hemin, CatG4, and either streptavidin or biotin tyramide in the absence of H_2O_2 is enough to support what appear to be streptavidin shifted products. It is also of note that the streptavidin shifted products in lane 6, formed in the absence of biotin tyramide but with H_2O_2, have a distinc mobility from the streptavidin shifted products formed in the absence of H_2O_2 (lane 3, 7, 8). This change in mobility between products, best observed in lanes 3 and 6 where biotin tyramide is not present, may mean a different mode of attachment and/or preferred reactive site for the two reactions. NMR analyses of these complexes may be valuable to elucidate this.

A comparison between lane 2 and 5 highlights the different biotin tyramide products created without and with H_2O_2. In the absence of H_2O_2 (lane 2), the shifted products represent 5% of the lane counts and form a relatively well resolved band. With H_2O_2 (lane 5) a large smear appears. Quantification of the radioactive counts using software identified that the smear represents 58% of the lane counts, indicating the peroxide supported reaction creates a larger range of biotinylated products.

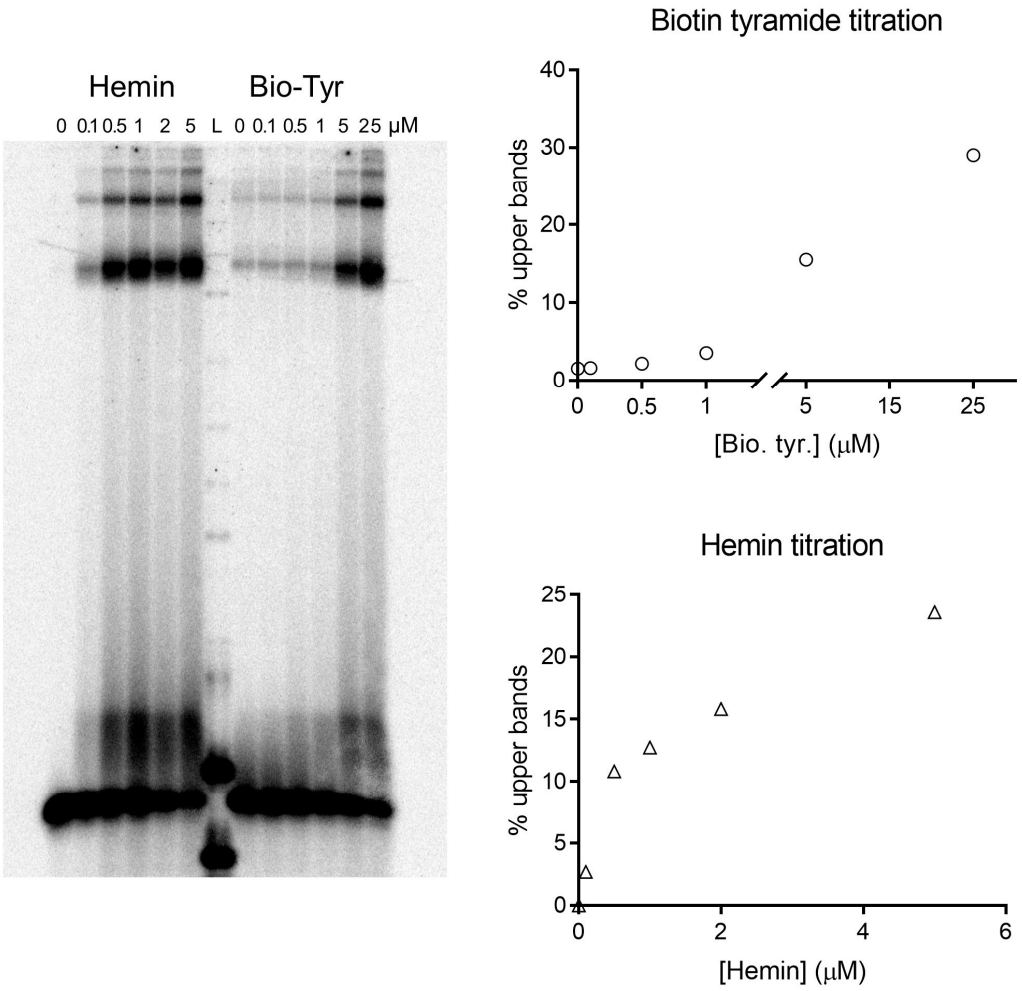


Figure 4-5 Heme-dependence and biotin-tyramide dependence for G4 biotinylation. 10% denaturing gel. The standard condition was 1 µM CatG4, 1 µM hemin, 5 µM biotin tyramide and 250 µM H_2O_2, with each lane modifying a reagent

4.4.2. Optimization

To characterize the efficiency of the reaction and to determine the ideal concentrations of the various reactants, the following experiments were done. Figure 4-5 shows the biotin tyramide and hemin dependency of the reaction. Modifications to hemin or to biotin tyramide concentration were made independently, leaving all other conditions the same. The percentage of bands shifted by streptavidin was used as the measure for efficacy. Hemin titration has a large influence at low concentrations, whereas biotin tyramide begins to show an effect above background levels at concentrations at 5 µM and

above. Below these higher concentrations, it is likely that the local concentration of tyramide molecules is too low for appreciable activation.

A time course for the tyramide assay was then done for CatG4, under the conditions of 1 μM CatG4, 1 μM hemin, 5 μM biotin tyramide and 250 μM H_2O_2, which were deemed to be the minimal and best conditions from the previous experiment. The reaction was carried out for 30 minutes. To avoid side reactions, aliquots were taken throughout this period and any remaining H_2O_2 was quenched by the addition of catalase enzyme. The first time point of 0.1 minutes represents the time taken, immediately after the addition of H_2O_2, to transfer an aliquot and add catalase. The results are displayed in Figure 4-6, and show that the reaction products plateau after the first 5 minutes, as calculated by a quantification of the percentage of CatG4 shifted by streptavidin over the time period.

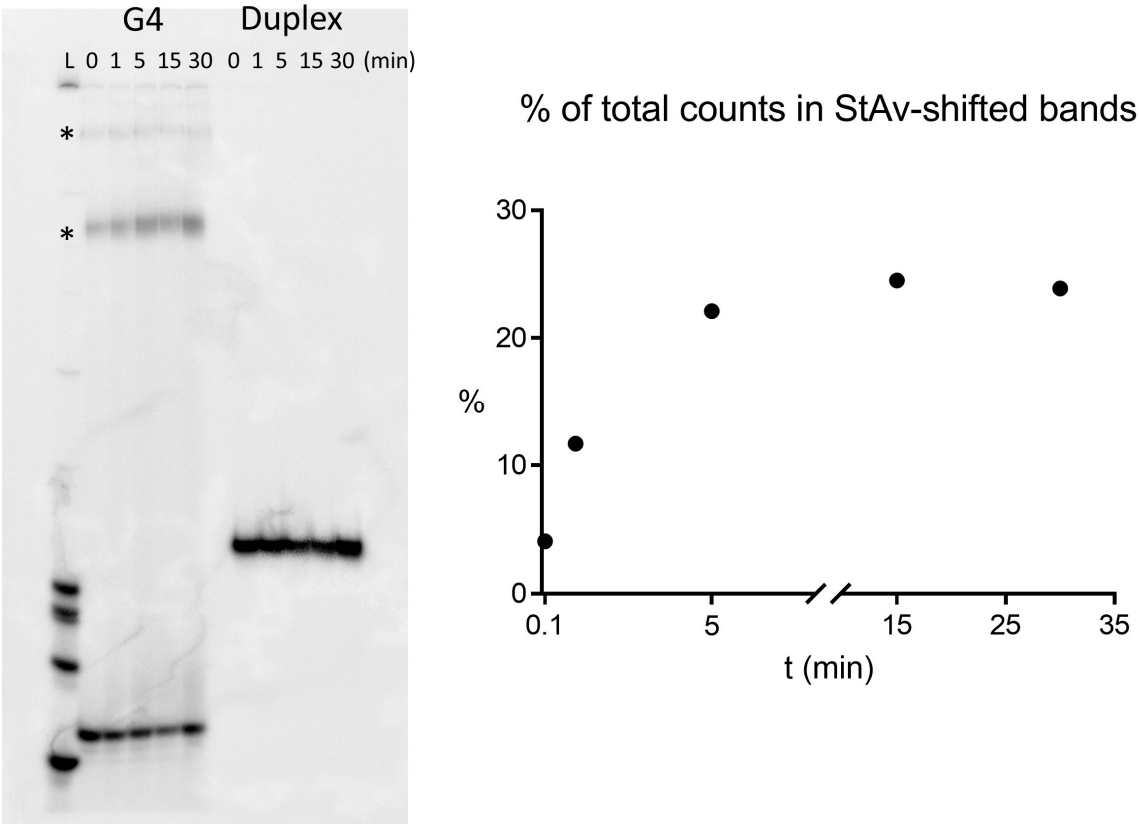


Figure 4-6 **Time course of biotin tyramide assay. Samples were quenched with catalase enzyme at each time point. 10% native gel. 1 μM CatG4, 1 μM hemin, 5 μM biotin tyramide and 250 μM H_2O_2.**

4.4.3. Effect of non-quadruplex DNA

The following experiments were done to test the viability of this assay in the presence of duplex DNA, as well as whether duplex DNA would be biotinylated if present during the reaction. Figure 4-7 demonstrates the result of the biotin tyramide assay on CatG4 in the presence of an excess of unlabeled duplex DNA. The addition of 10 times (lane 3) and 100 times (lane 4) molar excess of a 28 base pair duplex to the standard tyramide assay procedure does not have a significant effect on the percent of CatG4 that is biotinylated.

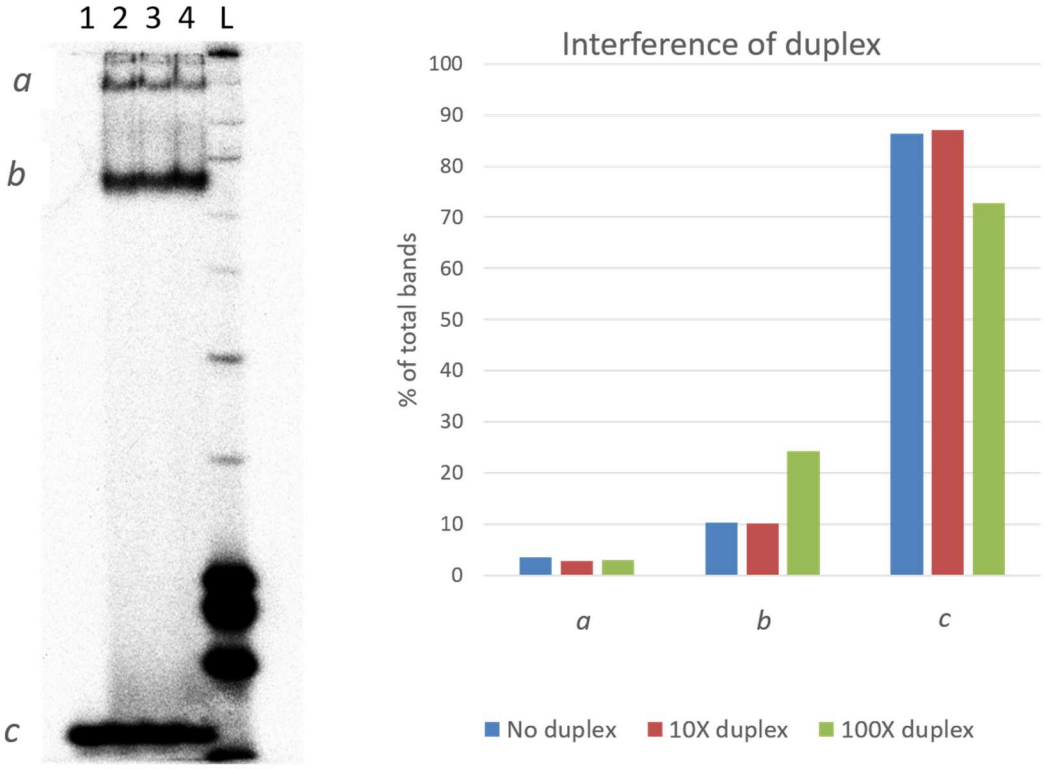


Figure 4-7 Interference of duplex DNA on the biotin tyramide assay. 10% native gel.

A similar experiment was conducted by radiolabeling both duplex and CatG4 and subjecting the mixture to biotin tyramide treatment, shown in Figure 4-8. CatG4 (G4) and duplex (D) samples were treated with the biotin tyramide assay individually and in 1:1 μM and 10:1 μM duplex/G4 ratios. The streptavidin shifting pattern of CatG4 (arrows) remains the same whether it was treated in isolation (lane 8), treated in a 1:1 mixture (lane 12), or a 10:1 mixture (lane 14). Similarly, the duplex band is unchanged in the same three scenarios (lanes 9, 12, 14), suggesting that there is little to no tagging of non-quadruplex

DNA. It would be useful to look at the assays performance in much more crowded conditions to closer mimic intracellular conditions. The specificity of the reaction bodes well for *in vivo* applications.

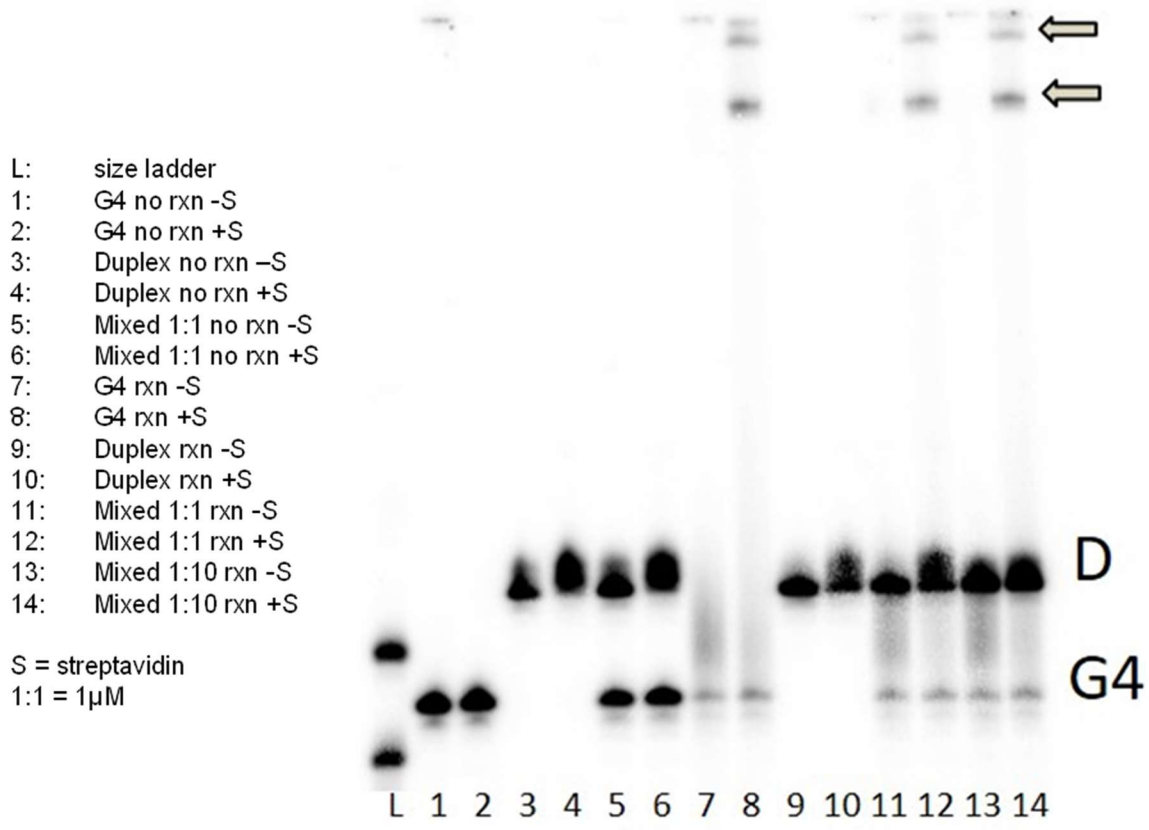


Figure 4-8 Treatment of duplex and quadruplex DNA with biotin tyramide assay together and seperately. Radioactively labeled duplex DNA does not change mobility, while CatG4 (G4) is shifted by streptavidin. 10% denaturing gel.

Sequence Name	*Bases*	*Sequence*
CG4- 5'-extended	40	5' - TAG TTG AAT GAC TTG ATT TTG GGT AGG GCG GGT TGG GAA A
CG4- 5'-extended complementary strand	16	5' - TCA AGT CAT TCA ACT A
CG4- 3'-extended	40	5' - TGG GTA GGG CGG GTT GGG AAA TTT AGT TCA GTA AGT TGA T
CG4- 3'-extended complementary strand	16	5' - ATC AAC TTA CTG AAC T
CG4 (pure G-quadruplex)	21	5' - TGG GTA GGG CGG GTT GGG AAA

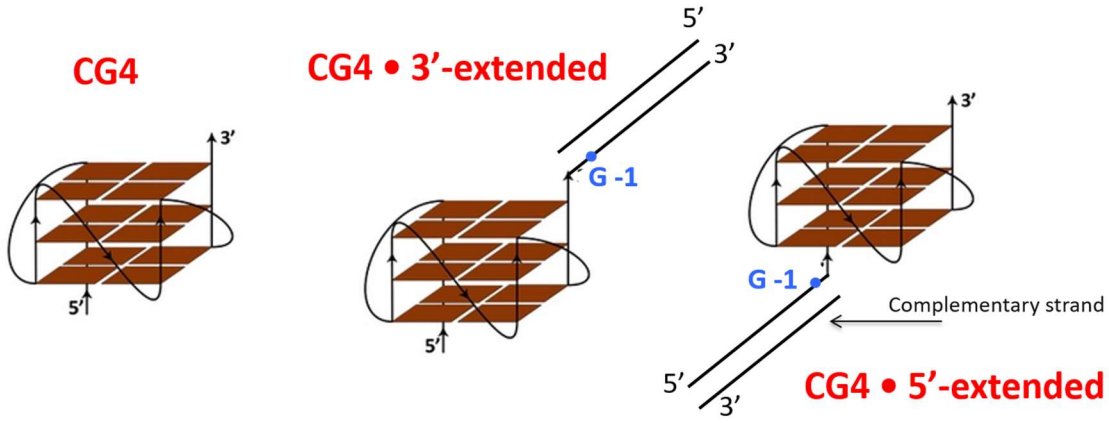


Figure 4-9 Sequences designed to test the distance dependency of the the tyramide radical.

4.4.4. Distance dependency of reaction

As we saw in Chapter 3, the treatment of damaged or modified DNA with alkali conditions is known to induce cleavage of the phosphate backbone at the site of modification. These alkali labile lesions are most often alkylations of the N7 position, specifically on guanine, but other modifications can be cleaved, such as the oxidized guanine bases modified during the peroxidase reaction. With the knowledge that the tyramide radical should target the bases of DNA, it could be possible to footprint these locations with an alkali cleavage protocol, such as piperidine footprinting. This procedure was performed on CatG4 in an attempt to understand the most common modification sites, if multiple biotinylation events were occurring, and the distance that the tyramide radical could migrate. For these experiments, modifications were made to the CatG4 sequence (Figure 4-9). The 3' and 5' ends of the quadruplex were extended with unrelated sequence, to serve as potential distance markers. The complementary sequences to these extended regions were also obtained to increase their rigidity.

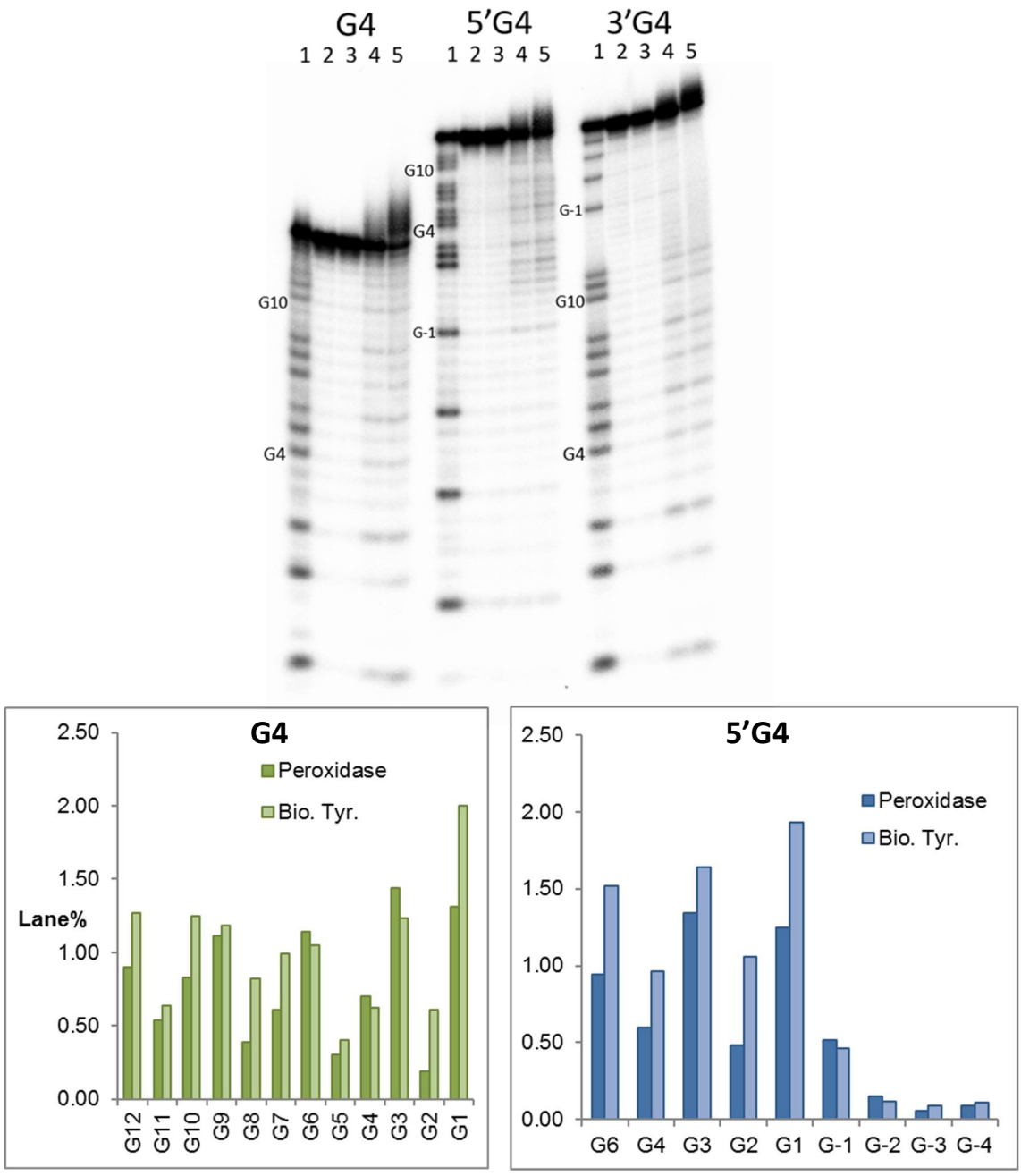


Figure 4-10 Piperidine treatment of peroxide reaction vs. biotin tyramide assay. Lane contents: 1-DNA alone, 2 DNA-hemin, 3-peroxidase reaction, 4-biotin tyramide assay. Reactions contain 1 µM DNA, 1 µM hemin, 5 µM biotin tyramide and 250 µM H_2O_2.

Due to the nature of the peroxidase and tyramide reactions, they both are likely to target predominantly guanine bases. To determine if the products of these two reactions could be differentiated, a comparison of the response to piperidine treatment was made on CatG4 and the two extended versions (Figure 4-10). The similarity between the

peroxidase footprint (lanes 4, all constructs) and the biotin tyramide footprint (lanes 5, all constructs) is evident. Quantification of the bands intensities as a percentage of the entire lane intensity identifies minor differences between tyramide and peroxide treatment. Tyramide treatment shows generally higher levels of damage at guanines in the quadruplex.

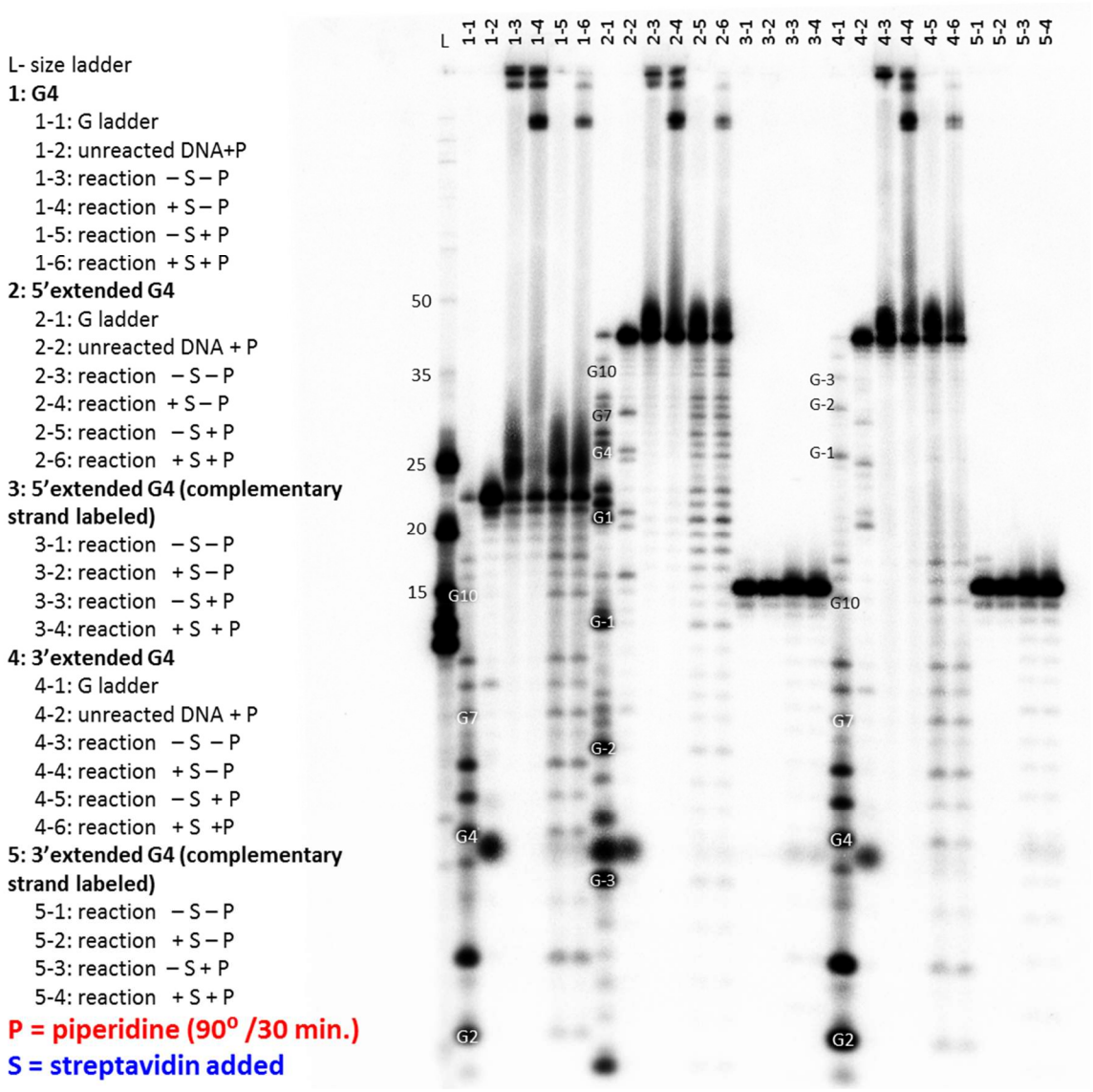


Figure 4-11 Comparison of piperidine and non piperidine treated damage patterns with and without streptavidin. 12% denaturing gel.

To examine whether tyramide damage patterns could be further differentiated from the peroxide footprint, biotin tyramide treated samples of the CatG4 variants in Figure 4-9 were ethanol precipitated, treated with piperidine, and then compared on a gel with and without streptavidin (Figure 4-11). Quantification of the banding pattern for G4 can be seen in Figure 4-12, and demonstrates a lack of any distinguishable features between streptavidin shifted and unshifted piperidine foot printing. The modified sequences are likely to have changed in mobility so they no longer represent a given band on the gel. In Figure 4-11, the streptavidin shifted species are seen to decrease once the sample is piperidine treated, meaning that base treatment is affecting the biotin tyramide modification in a way that hinders the ability for streptavidin to bind. However, the majority of the streptavidin product that no longer binds still remains in the biotin smear, which may indicate that the modification damages the biotin moiety specifically.

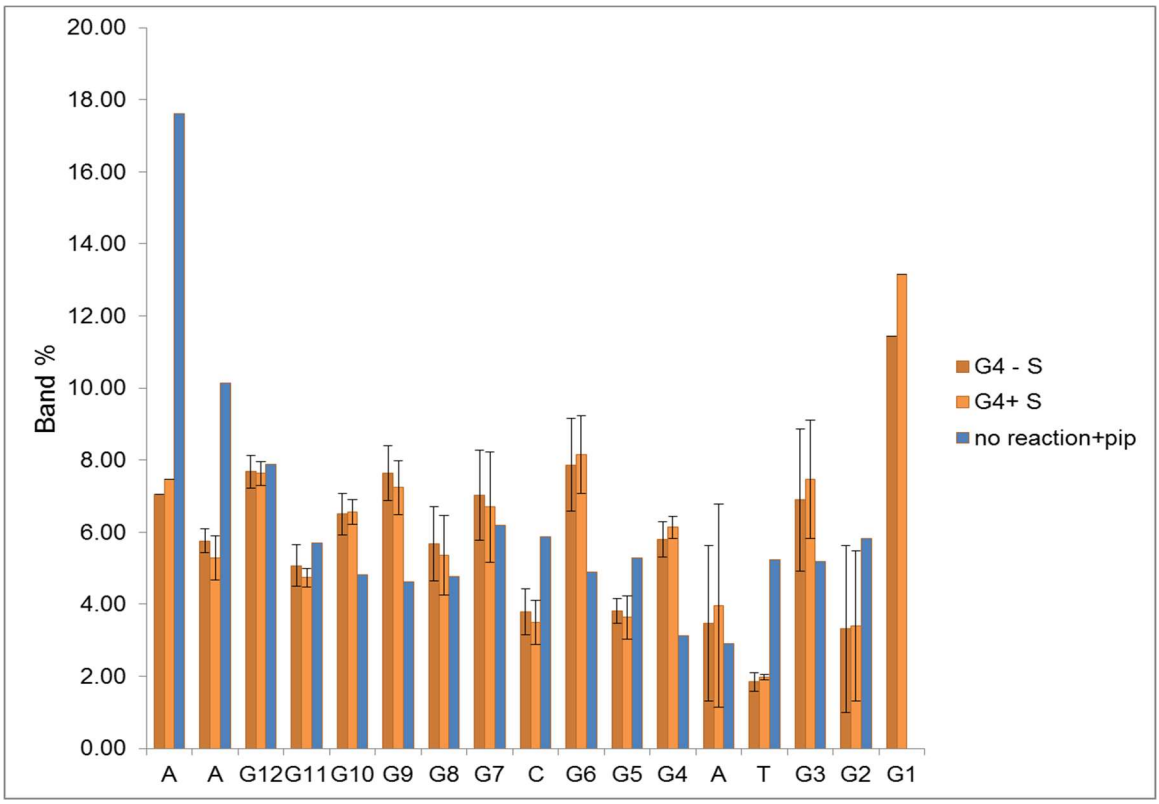


Figure 4-12 Quantification of damage from biotin tyramide assay treated with piperidine (pip) by nucleotide of CatG4 (G4). A comparison is made with and without streptavidin. Data from Figure 4-11. Error bars are s.d. from 2 replicates.

	Tyramide rxn	Fenton rxn (mins)	Streptavidin
controls			
1.1	-	-	-
1.2	-	-	+
1.3	+	-	-
1.4	+	-	+
G ladder			
2.1	-	1	-
2.2	+	1	-
2.3	-	1	+
2.4	+	1	+
3.1	-	2	-
3.2	+	2	-
3.3	-	2	+
3.4	+	2	+
4.1	-	3	-
4.2	+	3	-
4.3	-	3	+
4.4	+	3	+
5.1	-	4	-
5.2	+	4	-
5.3	-	4	+
5.4	+	4	+
6.1	-	10	-
6.2	+	10	-
6.3	-	10	+
6.4	+	10	+

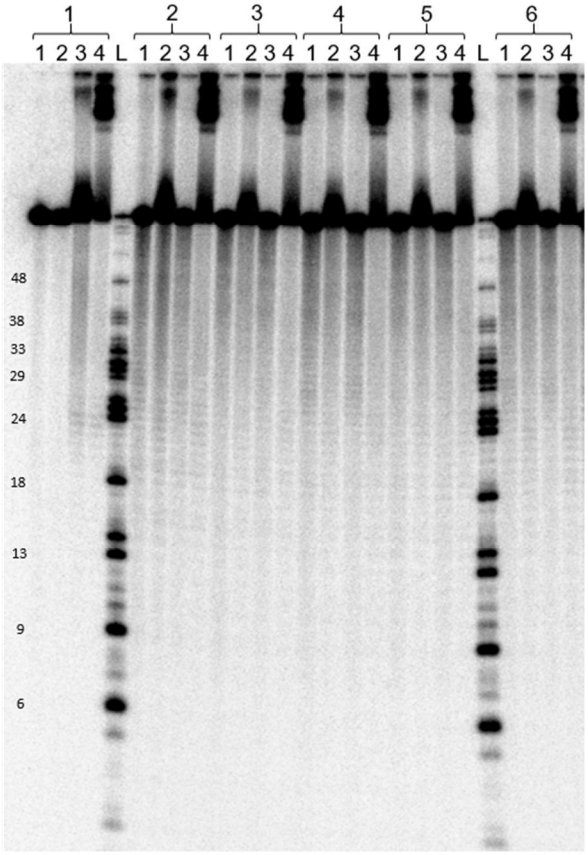


Figure 4-13 Fenton reaction on an extended version of CatG4. 12% denaturing gel. Fenton G4 DNA was treated with 1 μM CatG4, 1 μM hemin, 5 μM biotin tyramide and 250 μM H_2O_2 prior to administeration of the Fenton conditions.

We attempted to use the Fenton reaction for further information on the localization of the biotin tag on CatG4. Using Fe(II) to generate hydroxyl radical based cleavage should yield a footprint unbiased by the peroxidase reaction. Treating this footprint with streptavidin could then identify shifting patterns based upon areas of high modification. The result is shown in Figure 4-13 and shows time points of a Fenton reaction on biotin tyramide labelled DNA. The time points gave similar results, indicating that more hydroxyl radical generation was needed. While it is difficult to make any precise analysis, there is a noticeable reduction in smearing of fragments below the full-length DNA when samples are treated with piperidine. A comparison of lanes 2-2 and 2-4 highlights the streptavidin induced relocation of many of the Fenton generated products to mobilities above that of the full-length DNA. The same phenomenon can be seen in the absence of the Fenton

reaction in lanes 1-3 and 1-4, where the damage induced by the tyramide reaction can be shifted with streptavidin. This mobility shift appears to occur for fragments beginning roughly at the G-quadruplex core at G24, supporting the evidence that tyramide has a preference for guanine bases and acts at a very short distance.

4.4.5. PCR

In order to carry out a successful *in vivo* pulldown of G-quadruplex forming structures, it would be advantageous to be able to amplify the sequences, as they would likely be present in low abundance. The plausibility of using PCR to amplify biotin tagged sequences was investigated through the design of a 3' and 5' extended version of CatG4 (Figure 4-14), with the ends of these extensions serving as primers. The reverse, template binding primer is a nested primer and was used to ensure the product was a shorter size than the template.

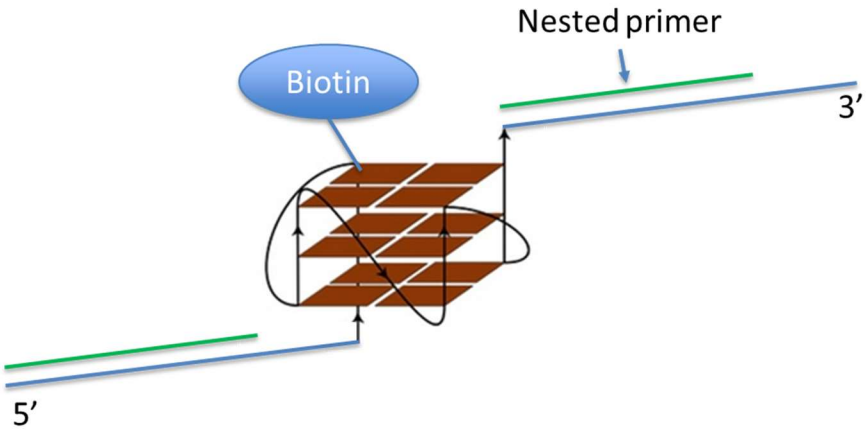


Figure 4-14 Design of the CatG4 variant for the PCR reactions. The 3' end contains a nested primer to differentiate from the full length template.

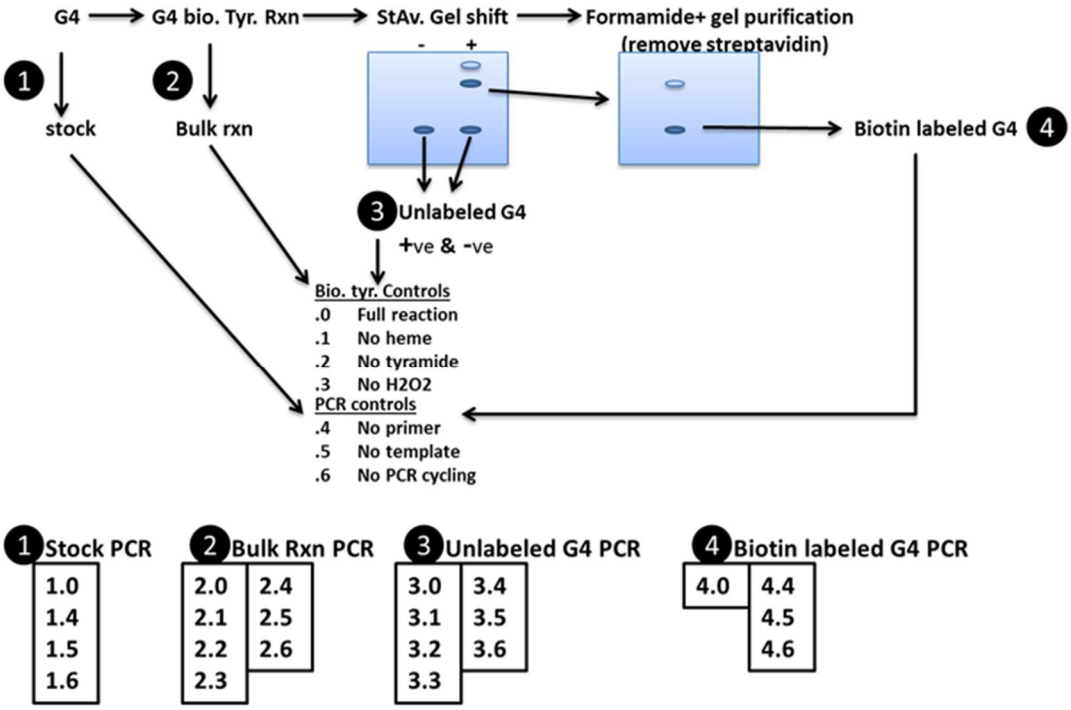
Figure 4-15 (a and b) highlights how the purification of biotin tyramide labeled sequences was done. The general protocol was to react the quadruplex template with biotin tyramide and gel purify the streptavidin shifted portion as normally, but then remove the streptavidin from the biotinylated template through formamide, EDTA and heating, followed by another denaturing gel. This is predominantly biotinylated template as can be seen in comparing the template bands from lane 1.0 and lane 4.6, where the entirety of the template band is lower in mobility. These purified labelled templates serve as the experimental condition for whether or not Taq polymerase is able to proceed past these

modifications. The biotin tyramide assay reaction conditions were changed to increase the overall proportion of tagged sequences- 1 µm DNA, 5 µM hemin, 1 mM H_2O_2 and either low (L) 20 µM or high (H) 200 µM biotin tyramide.

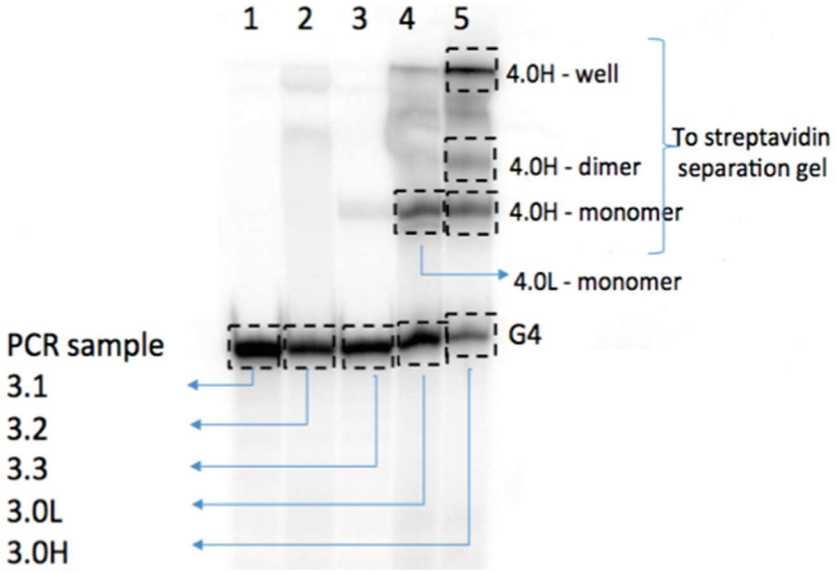
The results of the PCR reaction are seen in Figure 4-15 (c), and show a successful amplification of the biotin tyramide labelled sequence in lanes beginning with 4. The arrow on the right side highlights the 84 nt template band of DNA, and the star is the expected 74 nt PCR product. The template band is smeared in the biotinylated group (4) and resolved in the control untreated group (1) and the unshifted group (3), which suggests the majority of template is labelled. The amount of PCR product is relatively constant across all templates. Comparing the PCR product from 1.0, 4.0H-mono, 4.0H-dimer, and 4.0H-well ('mono' signifies the most mobile of the streptavidin shifted products, 'dimer' is the band above that, and 'well' is product recovered from the well of the gel), there no difference in the quantity of the product. That there is no change from 0 biotinylations to possible multiple biotinylations suggests PCR amplification is compatible with this assay, even in extreme conditions. Lane 1.5 indicates that there are some primer products being created in the absence of template. A closer look at amplification of the stock (1.0) identifies almost complete consumption of the primer, which is not the case for any of the 4.X labeled conditions.

The design of the controls that are categorized as X.4 act similar to a primer extension assay; they lack forward primer, with extension only occurring from the reverse primer annealed to the template. This should give a direct indication of sites where extension was halted. In lane 1.4 we can see a region prior to the 40 nt ladder marker where extension stopped. It is marked with a black arrow and is consistent across most of the PCR samples, and is likely due to G-quadruplex formation. From the 3' end of the template, the first guanine in the G-quadruplex begins at 47 nts. After this point, there are several more stop sites between the 40 and 60 nt regions which appear in groups of 3 nt. These are likely representative of the other guanine regions in the quadruplex. A comparison of 1.4 with the "primer extension" of the bulk reaction in 2.4 identifies a similar region for extension halting. Lane 2.4 has more, and more general extension halting, which supports the premise of tyramide labeling favoring the quadruplex. The pattern does not implicate any particular favoured position for tyramide attack.

a


G4
G4 bio. Tyr. Rxn
StAv. Gel shift
Formamide+ gel purification
(remove streptavidin)
- +
1
stock
2
Bulk rxn
Biotin labeled G4
4
3 Unlabeled G4
+ve & -ve
Bio. tyr. Controls
.0 Full reaction
.1 No heme
.2 No tyramide
.3 No H2O2
PCR controls
.4 No primer
.5 No template
.6 No PCR cycling
1 Stock PCR
1.0
1.4
1.5
1.6
2 Bulk Rxn PCR
2.0 2.4
2.1 2.5
2.2 2.6
2.3
3 Unlabeled G4 PCR
3.0 3.4
3.1 3.5
3.2 3.6
3.3
4 Biotin labeled G4 PCR
4.0 4.4
4.5
4.6

b


1 2 3 4 5
4.0H - well
4.0H - dimer
4.0H - monomer
To streptavidin separation gel
4.0L - monomer
G4
PCR sample
3.1
3.2
3.3
3.0L
3.0H

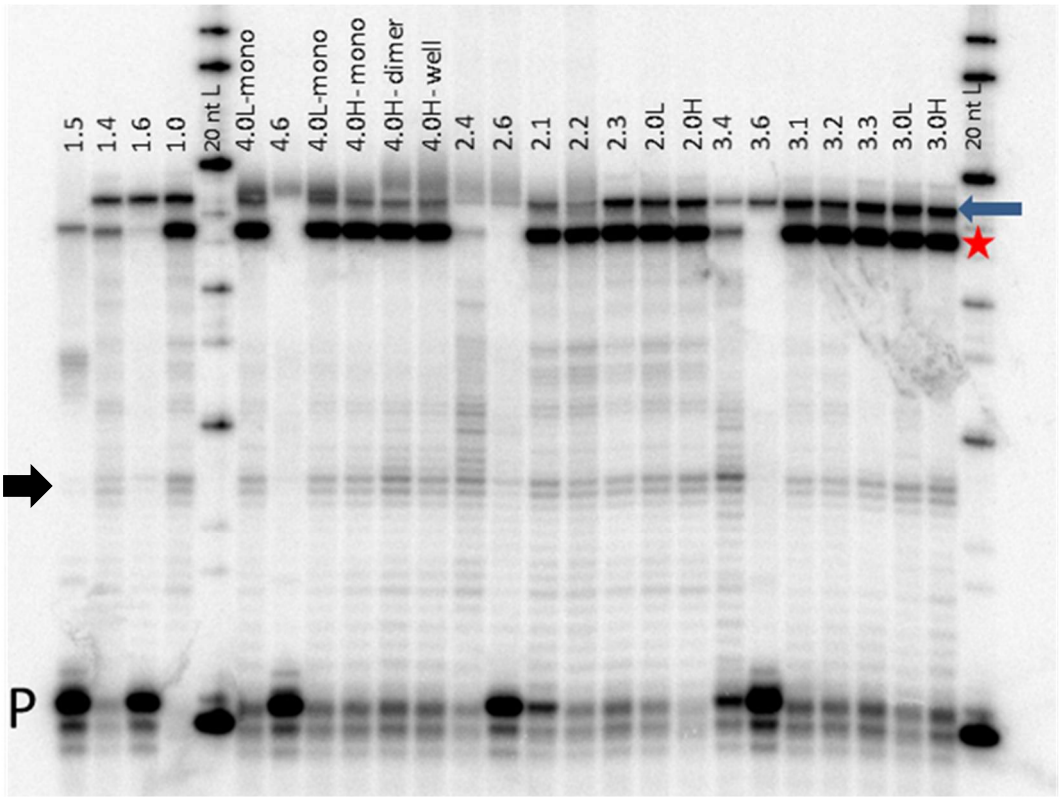


Figure 4-15 **Investigation of whether bio-tyr mediated covalent biotinylation of G4 DNA enables such DNA to be amplified by PCR. *a*: A scheme for treatment of G4 with bio-tyr to generate biotinylated DNAs and also various unbiotinylated controls. In block (1) G4 "stock" refers to G4 DNA that has not been subjected to a bio-tyr-mediated biotinylation reaction. Block (2) "bulk reaction" refers to G4 DNA that has subjected to the bio-tyr reaction, but not subsequently subjected to a purification protocol to separate out biotinylated from unbiotinylated DNA. "L" and "H" refer to low (20 μM) and high (200 μM) concentrations of bio-tyr, respectively. *b*: A further elaboration on blocks (3) and (4). In this preparatory gel, lane 1 shows bio-tyr reaction conditions but excluding hemin; lane 2 shows the reaction excluding bio-tyr; lane 3 excludes H_2O_2; lane 4 shows the full reaction with all components, including 20 μM bio-tyr; lane 5 shows full reaction but with 200 μM bio-tyr. *c*: Gel showing the results of experiment schematized in *a* and *b*. The blue arrow indicates the lightly ^{32}P-labeled template (84 nt) whereas the red star indicates the PCR-amplified product (74 nt). The black arrow indicates a common interruption point. "P" refers to the ^{32}P-labeled reverse primer. "20 nt" shows a reference ladder.**

4.5. Conclusion

The results here show a way that the peroxidase activity of G-quadruplex-hemin complexes can be exploited to self-biotinylate. The reaction proceeds with good efficiency *in vitro* and is specific in labelling for G-quadruplexes and not duplex DNA. These biotin tagged sequences have been shown to be compatible with PCR amplification; an important factor to consider if we were to this for *in vivo* pulldown of G-quadruplexes as they would likely be in low abundance.

With some characteristics of this reaction now elucidated *in vitro*, the question arises as to its usefulness *in vivo*. This reaction may be able to address the question of whether or not heme binds and is activated by G-quadruplexes in the cell in either homeostatic or disease conditions. One determining factor for the success of this would be the abundance of heme that is available in the cell. “Free” heme (unbound to protein) is tightly regulated and is likely present in concentrations <1 μM in most cell compartments.[78] Given the low K_d’s of many G-quadruplexes, including the hybrids discussed in Chapter 3, it could be possible that binding may occur at low hemin concentrations. If “free” heme in the cell is not of sufficient concentration to activate the reaction in a detectable manner, there are ways to supplement cell growth media with heme to allow uptake and increase levels. This has been used previously to activate heme enzymes from poor to high catalytic activity.[79,80] This may help achieve higher levels of hemin-quadruplex complexation, although at the expense of biological relevance.

This technique could also be used for the visualization or pulldown of intracellular G-quadruplexes. If quadruplexes are able to be tagged by the reactive substrate, it would be straightforward to isolate them from cell lysate with streptavidin and do further analyses. Further experiments *in vitro* may be useful to determine if biotin tyramide radicals are capable of tagging nearby proteins as it may also be possible to do proteomic screens of quadruplex interacting proteins. MS analyses is highly compatible with this tyramide tagging system, as demonstrate by Ting and coworkers.[73] The topologies of quadruplexes that are capable of being pulled down will be important to investigate further. While the duplex-quadruplex hybrid structures bind, and activate hemin, Construct 1 was unable to oxidize biotin tyramide substrate. It would be valuable to see if other reactive substrates may be more compatible with that type of quadruplex morphology. Increasing

the breadth of quadruplex morphologies that are compatible would only increase its applicability.

References

1. Gregory, T. R. Synergy between sequence and size in Large-scale genomics. *Nat. Rev. Genet.* **6,** 699–708 (2005).

2. Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34,** 564–574 (2006).

3. Kruger, K. *et al.* Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31,** 147–157 (1982).

4. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35,** 849–857 (1983).

5. Hurst, H. C. Transcription factors. 1: bZIP proteins. *Protein Profile* **1,** 123–168 (1994).

6. Menkens, A. E., Schindler, U. & Cashmore, A. R. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem. Sci.* **20,** 506–510 (1995).

7. Shimizu, H. *et al.* LIP19, a basic region leucine zipper protein, is a Fos-like molecular switch in the cold signaling of rice plants. *Plant Cell Physiol.* **46,** 1623–1634 (2005).

8. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249,** 505–510 (1990).

9. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).

10. Ellington, A. D. & Szostak, J. W. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355,** 850–852 (1992).

11. Wu, X. *et al.* Cell-SELEX Aptamer for Highly Specific Radionuclide Molecular Imaging of Glioblastoma In Vivo. *PLOS ONE* **9,** e90752 (2014).

12. Huizenga, D. E. & Szostak, J. W. A DNA Aptamer That Binds Adenosine and ATP. *Biochemistry (Mosc.)* **34,** 656–665 (1995).

13. Yang, K.-A. *et al.* Recognition and sensing of low-epitope targets via ternary complexes with oligonucleotides and synthetic receptors. *Nat. Chem.* **6,** 1003–1008 (2014).

14. Jo, M. *et al.* Development of Single-Stranded DNA Aptamers for Specific Bisphenol A Detection. *Oligonucleotides* **21,** 85 (2011).

15. Breaker, R. R. & Joyce, G. F. A DNA enzyme that cleaves RNA. *Chem. Biol.* **1,** 223–229 (1994).

16. Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 4262–4266 (1997).

17. Santoro, S. W. & Joyce, G. F. Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry (Mosc.)* **37,** 13330–13342 (1998).

18. Li, J., Zheng, W., Kwon, A. H. & Lu, Y. In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. *Nucleic Acids Res.* **28,** 481–488 (2000).

19. Flynn-Charlebois, A. *et al.* Deoxyribozymes with 2'-5' RNA ligase activity. *J. Am. Chem. Soc.* **125,** 2444–2454 (2003).

20. Chandra, M. & Silverman, S. K. DNA and RNA Can Be Equally Efficient Catalysts for Carbon−Carbon Bond Formation. *J. Am. Chem. Soc.* **130,** 2936–2937 (2008).

21. Chinnapen, D. J.-F. & Sen, D. A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 65–69 (2004).

22. Travascio, P., Bennet, A. J., Wang, D. Y. & Sen, D. A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites. *Chem. Biol.* **6,** 779–787 (1999).

23. Li, Y. & Sen, D. A catalytic DNA for porphyrin metallation. *Nat. Struct. Mol. Biol.* **3,** 743–747 (1996).

24. Cochran, A. G. & Schultz, P. G. Antibody-catalyzed porphyrin metallation. *Science* **249,** 781–783 (1990).

25. Ralph, R. K., Connors, W. J. & Khorana, H. G. Secondary Structure and Aggregation in Deoxyguanosine Oligonucleotides. *J. Am. Chem. Soc.* **84,** 2265–2266 (1962).

26. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. Proc. Natl. Acad. Sci. U. S. A. 48, 2013–2018 (1962).

27. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334,** 364–366 (1988).

28. Sen, D. & Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344,** 410–414 (1990).

29. Bhattacharyya, D., Mirihana Arachchilage, G. & Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **4,** (2016).

30. Venczel, E. A. & Sen, D. Parallel and antiparallel G-DNA structures from a complex telomeric sequence. *Biochemistry (Mosc.)* **32,** 6220–6228 (1993).

31. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13,** 770–780 (2012).

32. Schaffitzel, C. *et al.* In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl. Acad. Sci.* **98,** 8572–8577 (2001).

33. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5,** 182–186 (2013).

34. Smith, J. S. *et al.* Rudimentary G-quadruplex–based telomere capping in Saccharomyces cerevisiae. *Nat. Struct. Mol. Biol.* **18,** 478–485 (2011).

35. Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* **3,** 218–221 (2007).

36. Gomez, D. *et al.* Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.* **32,** 371–379 (2004).

37. Zaug, A. J., Podell, E. R. & Cech, T. R. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 10864–10869 (2005).

38. Schaffitzel, C., Postberg, J., Paeschke, K. & Lipps, H. J. Probing telomeric G-quadruplex DNA structures in cells with in vitro generated single-chain antibody fragments. *Methods Mol. Biol. Clifton NJ* **608,** 159–181 (2010).

39. Travascio, P., Li, Y. & Sen, D. DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chem. Biol.* **5,** 505–517 (1998).

40. Franzen, S. & Boxer, S. G. On the Origin of Heme Absorption Band Shifts and Associated Protein Structural Relaxation in Myoglobin following Flash Photolysis. *J. Biol. Chem.* **272,** 9655–9660 (1997).

41. Sen, D. & Poon, L. C. H. RNA and DNA complexes with hemin [Fe(III) heme] are efficient peroxidases and peroxygenases: how do they do it and what does it mean? *Crit. Rev. Biochem. Mol. Biol.* **46,** 478–492 (2011).

42. Saito, K., Tai, H., Hemmi, H., Kobayashi, N. & Yamamoto, Y. Interaction between the heme and a G-quartet in a heme-DNA complex. *Inorg. Chem.* **51,** 8168–8176 (2012).

43. Travascio, P., Witting, P. K., Mauk, A. G. & Sen, D. The Peroxidase Activity of a Hemin−DNA Oligonucleotide Complex: Free Radical Damage to Specific Guanine Bases of the DNA. *J. Am. Chem. Soc.* **123,** 1337–1348 (2001).

44. Travascio, P., Sen, D. & Bennet, A. J. DNA and RNA enzymes with peroxidase activity - An investigation into the mechanism of action. *Can. J. Chem.* **84,** 613–619 (2006).

45. Rojas, A. M., Gonzalez, P. A., Antipov, E. & Klibanov, A. M. Specificity of a DNA-based (DNAzyme) peroxidative biocatalyst. *Biotechnol. Lett.* **29,** 227–232 (2007).

46. Shumayrikh, N., Huang, Y. C. & Sen, D. Heme activation by DNA: isoguanine pentaplexes, but not quadruplexes, bind heme and enhance its oxidative activity. *Nucleic Acids Res.* **43,** 4191–4201 (2015).

47. Lai, R. Y., Plaxco, K. W. & Heeger, A. J. Aptamer-Based Electrochemical Detection of Picomolar Platelet-Derived Growth Factor Directly in Blood Serum. *Anal. Chem.* **79,** 229–233 (2007).

48. Zuo, X. *et al.* A Target-Responsive Electrochemical Aptamer Switch (TREAS) for Reagentless Detection of Nanomolar ATP. *J. Am. Chem. Soc.* **129,** 1042–1043 (2007).

49. Tang, Y., Ge, B., Sen, D. & Yu, H.-Z. Functional DNA switches: rational design and electrochemical signaling. *Chem. Soc. Rev.* **43,** 518–529 (2013).

50. Fahlman, R. P. & Sen, D. DNA conformational switches as sensitive electronic sensors of analytes. *J. Am. Chem. Soc.* **124,** 4610–4616 (2002).

51. Thomas, J. M., Chakraborty, B., Sen, D. & Yu, H.-Z. Analyte-Driven Switching of DNA Charge Transport: De Novo Creation of Electronic Sensors for an Early Lung Cancer Biomarker. *J. Am. Chem. Soc.* **134,** 13823–13833 (2012).

52. Thomas, J. M., Yu, H.-Z. & Sen, D. A Mechano-Electronic DNA Switch. *J. Am. Chem. Soc.* **134,** 13738–13748 (2012).

53. Clever, G. H., Kaul, C. & Carell, T. DNA--metal base pairs. *Angew. Chem. Int. Ed Engl.* **46,** 6226–6236 (2007).

54. Liu, B., Li, D. & Shang, H. General peroxidase activity of a parallel G-quadruplex-hemin DNAzyme formed by Pu39WT - a mixed G-quadruplex forming sequence in the Bcl-2 P1 promoter. *Chem. Cent. J.* **8,** 43 (2014).

55. Cheng, X., Liu, X., Bing, T., Cao, Z. & Shangguan, D. General Peroxidase Activity of G-Quadruplex−Hemin Complexes and Its Application in Ligand Screening. *Biochemistry (Mosc.)* **48,** 7817–7823 (2009).

56. Grigg, J. C., Shumayrikh, N. & Sen, D. G-Quadruplex Structures Formed by Expanded Hexanucleotide Repeat RNA and DNA from the Neurodegenerative Disease-Linked C9orf72 Gene Efficiently Sequester and Activate Heme. *PLOS ONE* **9,** e106449 (2014).

57. Chen, W. *et al.* DNA transducer-triggered signal switch for visual colorimetric bioanalysis. *Sci. Rep.* **5,** (2015).

58. Aleman-Garcia, M. A., Orbach, R. & Willner, I. Ion-Responsive Hemin–G-Quadruplexes for Switchable DNAzyme and Enzyme Functions. *Chem. – Eur. J.* **20,** 5619–5624 (2014).

59. Nakayama, S., Wang, J. & Sintim, H. O. DNA-Based Peroxidation Catalyst—What Is the Exact Role of Topology on Catalysis and Is There a Special Binding Site for Catalysis? *Chem. – Eur. J.* **17,** 5691–5698 (2011).

60. Murat, P. & Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* **25,** 22–29 (2014).

61. Lim, K. W., Khong, Z. J. & Phan, A. T. Thermal Stability of DNA Quadruplex–Duplex Hybrids. *Biochemistry (Mosc.)* **53,** 247–257 (2014).

62. Lim, K. W. & Phan, A. T. Structural Basis of DNA Quadruplex-Duplex Junction Formation. *Angew. Chem. Int. Ed.* **52,** 8566–8569 (2013).

63. Lim, K. W. *et al.* Duplex stem-loop-containing quadruplex motifs in the human genome: a combined genomic and structural study. *Nucleic Acids Res.* **43,** 5630–5646 (2015).

64. Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33,** 2908–2916 (2005).

65. Guédin, A., Gros, J., Alberti, P. & Mergny, J.-L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* **38,** 7858–7868 (2010).

66. Hänsel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* **48,** 1267–1272 (2016).

67. Paeschke, K., Capra, J. A. & Zakian, V. A. DNA replication through G-quadruplex motifs is promoted by the S. cerevisiae Pif1 DNA helicase. *Cell* **145,** 678–691 (2011).

68. Rodriguez, R. *et al.* Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **8,** 301–310 (2012).

69. Chambers, V. S. *et al.* High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33,** 877–881 (2015).

70. Faget, L. & Hnasko, T. S. Tyramide Signal Amplification for Immunofluorescent Enhancement. *Methods Mol. Biol. Clifton NJ* **1318,** 161–172 (2015).

71. Bobrow, M. N., Litt, G. J., Shaughnessy, K. J., Mayer, P. C. & Conlon, J. The use of catalyzed reporter deposition as a means of signal amplification in a variety of formats. *J. Immunol. Methods* **150,** 145–149 (1992).

72. Rhee, H.-W. *et al.* Proteomic Mapping of Mitochondria in Living Cells via Spatially-Restricted Enzymatic Tagging. *Science* **339,** 1328–1331 (2013).

73. Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11,** 456–475 (2016).

74. Speel, E. J., Hopman, A. H. & Komminoth, P. Amplification methods to increase the sensitivity of in situ hybridization: play card(s). *J. Histochem. Cytochem. Off. J. Histochem. Soc.* **47,** 281–288 (1999).

75. Mantle, P. G., Faucet-Marquis, V., Manderville, R. A., Squillaci, B. & Pfohl-Leszkowicz, A. Structures of Covalent Adducts between DNA and Ochratoxin A: A New Factor in Debate about Genotoxicity and Human Risk Assessment. *Chem. Res. Toxicol.* **23,** 89–98 (2010).

76. Dai, J., Wright, M. W. & Manderville, R. A. Ochratoxin a forms a carbon-bonded c8-deoxyguanosine nucleoside adduct: implications for c8 reactivity by a phenolic radical. *J. Am. Chem. Soc.* **125,** 3716–3717 (2003).

77. Kurzban, G. P., Bayer, E. A., Wilchek, M. & Horowitz, P. M. The quaternary structure of streptavidin in urea. *J. Biol. Chem.* **266,** 14470–14477 (1991).

78. Sassa, S. Why heme needs to be degraded to iron, biliverdin IXalpha, and carbon monoxide? *Antioxid. Redox Signal.* **6,** 819–824 (2004).

79. Richards, M. K. & Marletta, M. A. Characterization of Neuronal Nitric Oxide Synthase and a C415H Mutant, Purified from a Baculovirus Overexpression System. *Biochemistry (Mosc.)* **33,** 14723–14732 (1994).

80. Martell, J. D. *et al.* Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* **30,** 1143–1148 (2012).