<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>G-quadruplex DNA aptamers and their ligands: Structure, function and application</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Tucker, WO; Shum, KT; Tanner, JA</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Current Pharmaceutical Design, 2012, v. 18 n. 14, p. 2014-2026</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2012</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/146385">http://hdl.handle.net/10722/146385</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
G-quadruplex DNA Aptamers and their Ligands: Structure, Function and Application

W.O. Tucker, K.T. Shum and J.A. Tanner*

Department of Biochemistry, University of Hong Kong, Pokfulam, Hong Kong, China

Abstract: Highly specific and tight-binding nucleic acid aptamers have been selected against a variety of molecular targets for over 20 years. A significant proportion of these oligonucleotides display G-quadruplex structures, particularly for DNA aptamers, that enable molecular recognition of their ligands. G-quadruplex structures couple a common scaffold to varying loop motifs that act in target recognition. Here, we review DNA G-quadruplex aptamers and their ligands from a structural and functional perspective. We compare the diversity of DNA G-quadruplex aptamers selected against multiple ligand targets, and consider structure with a particular focus on dissecting the thrombin binding aptamer–thrombin interaction. Therapeutic and analytical applications of DNA G-quadruplex aptamers are also discussed. Understanding DNA G-quadruplex aptamers carries implications not only for therapeutics and diagnostics, but also in the natural biochemistry of guanine-rich nucleic acids.

Keywords: G-quadruplex, aptamer, thrombin, sclerostin, G-quartet, G4 tetrad, SELEX.

INTRODUCTION

The artificial selection of nucleic acid aptamers from a large pool of random sequences that bind specifically to a molecular target was first demonstrated contemporaneously by Tuerk & Gold and Ellington & Szostak in 1990 [1, 2]. Tuerk & Gold referred to the approach as systematic evolution of ligands by exponential enrichment (SELEX), whilst Ellington & Szostak first used the term ‘aptamer’ to describe the nucleic acids selected against a ligand in this manner Fig. (1). For both the original studies, selection was performed on RNA, which classically is assumed to have greater structural diversity than DNA due to the natural pucker in the ribose ring. The original pioneering studies selecting ligand-binding RNA were soon followed up with the selection of single-stranded DNA [3], demonstrating that DNA was able to bind targets in quite a similar manner to RNA. It was perhaps surprising the extent to which single-stranded DNA was able to demonstrate such structural plasticity, suggesting that DNA may also catalyze chemical transformations, and indeed the first deoxyribozyme was selected and characterized in 1994 [4]. The classic view of proteins ‘recognizing’ nucleic acids can be reframed in terms of aptamer selection; is it just as valid to consider nucleic acids ‘recognizing’ proteins? Single-stranded DNA can in ways be considered structurally and functionally similar to RNA, and that functionality depends on structure. One important structure often observed within single-stranded DNA aptamers is the G-quadruplex. Herein, we discuss the significance of the aptameric G-quadruplex and the roles it plays in molecular recognition by nucleic acid aptamers.

An aptamer-derived G-quadruplex was first observed in NMR studies on a thrombin-binding DNA aptamer [5], consisting of two tetrads of guanosine base pairs. This aptamer inhibited the thrombin-catalyzed polymerization of fibrinogen. It was shown that the G-quadruplex was critical for inhibitory activity as substitution of any of the guanines reduced the inhibition of thrombin [6]. Furthermore, the X-ray crystal structure of this aptamer in complex with thrombin was solved showing the G-quadruplex aptamer sandwiched between two positively charged regions of two separate thrombin molecules [7] Fig. (2). Remarkably, this remains the only structure of a G-quadruplex aptamer selected by SELEX in complex with its target. Since those original studies, the G-quadruplex...
### Table 1. DNA G-quadruplex Aptamers: Targets, Sequences and Topologies

<table>
<thead>
<tr>
<th>Target</th>
<th>Sequence (5’ to 3’)</th>
<th>Topology of G-Quadruplex</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein targeted DNA G-quadruplex Aptamers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thrombin</td>
<td>d(GGTTGGTGTGGTTGG)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[15]</td>
</tr>
<tr>
<td>Nucleolin</td>
<td>d(GGTTGGTGTGGTTGGTGG)</td>
<td>Anti-parallel G-quadruplex, dimer</td>
<td>[16]</td>
</tr>
<tr>
<td>Signal transducer and activator of transcription (STAT3)</td>
<td>d(GGGCGGGCGGGCGGCG)</td>
<td>N.A.</td>
<td>[17, 18]</td>
</tr>
<tr>
<td>Human RNase H1</td>
<td>d(CGGTCGCTCCGTGGTGGTTGGTGTCAGTA)</td>
<td>Anti-parallel G-quadruplex</td>
<td>[19]</td>
</tr>
<tr>
<td>Protein tyrosine phosphatase Shp2</td>
<td>d(AGCGTGAATACACACGCCGCGGTTGGGCTGGTGGCAGTGGC)</td>
<td>Parallel G-quadruplex</td>
<td>[20]</td>
</tr>
<tr>
<td>VEGF</td>
<td>d(TGTGGGGGACCGGGGTGGTAG)</td>
<td>Anti-parallel G-quadruplex</td>
<td>[21]</td>
</tr>
<tr>
<td>HIV-1 gp120 V3 loop</td>
<td>d(TTGGGTT)</td>
<td>Parallel G-quadruplex, tetramer</td>
<td>[22]</td>
</tr>
<tr>
<td>HIV-1 gp120 V3 loop / CD4 binding site</td>
<td>d(TGGGAG)</td>
<td>Parallel G-quadruplex</td>
<td>[23]</td>
</tr>
<tr>
<td>HIV-1 integrase</td>
<td>d(GGGTTGGGAGGAGGGT)</td>
<td>Anti-parallel G-quadruplex, dimer</td>
<td>[24]</td>
</tr>
<tr>
<td>HIV-1 reverse transcriptase</td>
<td>d(GGGGGTGGGAGGGTAGGCCTTAGGGTGGGGGGGCTAATGGAGCTCGTTGCTCAT)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[25]</td>
</tr>
<tr>
<td>HIV-1 reverse transcriptase</td>
<td>d(CGCCCTAGGATTAGCGACATTAGGCTGGGGGGGAGG)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[26]</td>
</tr>
<tr>
<td>HIV-1 nucleocapsid protein</td>
<td>d(GGGTTGGGAGGAGGGT)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[27]</td>
</tr>
<tr>
<td>HCV RNA dependent RNA polymerase</td>
<td>d(GGGCGGTGGTGTTGGGACTTAAATGTCGCTGTTT)</td>
<td>Anti-parallel G-quadruplex monomer</td>
<td>[28]</td>
</tr>
<tr>
<td>SARS –CoV helicase</td>
<td>d(AGCGGCGCATATGGTGTTGGGCTAGGATG)</td>
<td>Parallel G-quadruplex</td>
<td>[14]</td>
</tr>
<tr>
<td>M. tuberculosis polyphosphate kinase 2</td>
<td>d(AACACANGTGGTGGTGGGTGGGCTAGAAATAGG)</td>
<td>Anti-parallel G-quadruplex</td>
<td>[13]</td>
</tr>
<tr>
<td>Sclerostin</td>
<td>d(TTGGGTTGTAATGTGGGGGTTGGGCTGGTTG)</td>
<td>Parallel G-quadruplex</td>
<td>[12]</td>
</tr>
<tr>
<td>Insulin</td>
<td>d(GGGTTGGGAGGAGGGT)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[29]</td>
</tr>
<tr>
<td><strong>Non-protein targeted DNA G-quadruplex aptamers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hematoporphyrin IX</td>
<td>d(ATGGGTCGGCGCGGGCGG)</td>
<td>Anti-parallel G-quadruplex monomer</td>
<td>[30, 31]</td>
</tr>
<tr>
<td>Hemin</td>
<td>d(GGTGGTGGGTGGGCGG)</td>
<td>N.A.</td>
<td>[32]</td>
</tr>
<tr>
<td>Ochratoxin</td>
<td>d(GATCGGGTGGTGGCGTAAAAGGACATCGGACA)</td>
<td>N.A.</td>
<td>[33]</td>
</tr>
<tr>
<td>Potassium ions</td>
<td>d(GGTTAGGTTAGGGTAGG)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[34]</td>
</tr>
<tr>
<td>ATP</td>
<td>d(CCTGGGGGAGTATTGGGAGGAGAAGG)</td>
<td>Anti-parallel G-quadruplex, monomer</td>
<td>[35]</td>
</tr>
</tbody>
</table>

(NMR). Computationally, Quadruplex forming G-rich Sequences (QGRS) mapper is a web-based computer program that can analyse putative G-quadruplex forming sequences based on the primary sequence input [8].

G-quadruplex DNA aptamers have been observed after SELEX selection against a wide variety of targets – mainly proteins but also some other targets Table 1. The reader also interested in RNA G-quadruplexes is pointed towards two other recent reviews [9, 10]. One possible reason that G-quadruplex aptamers tend to be selected is the significant charge density of G-quadruplex DNA – twice the negative charge per unit length when compared to double-helical DNA [11]. The consequent electrostatic potential provides for strong binding to positively charged surfaces of proteins, although hydrophobic interactions are also critical for strong and specific
binding to the target. G-quadruplex structures also tend to be selected more often when performing selections using DNA libraries rather than RNA libraries, though this is perhaps unsurprising given RNA’s greater variety of structural topologies. G-quadruplex is one of only a few structural arrangements of DNA that involves sophisticated tertiary folding.

Common selection of DNA aptamers against totally unrelated targets has strengthened the hypothesis that G-quadruplex aptamers can target a broad range of protein structures, yet retain specificity. Indeed, we have recently isolated and characterized DNA aptamers against three different classes of protein targets --- SARS-CoV helicase, M. tuberculosis polyphosphate kinase 2 and sclerostin [12-14].

SARS-CoV helicase is a nucleic acid binding protein, polyphosphate kinase 2 is a polyanion binding protein and sclerostin is a protein interacting protein. We also employed the same selection strategy to isolate tight binding sequences using a magnetic bead approach so that there is no bias on selecting aptamers. Unexpectedly, G-quadruplexes were discovered to be present in all aptamers selected against these protein targets. Except for the G-quadruplex SARS-CoV helicase aptamers, all displayed inhibitory effects against their respective protein targets [12-14]. Taken together with other G-quadruplex forming aptamers in Table I, it can be surmised that the correlation between G-quadruplex forming aptamers and the class of protein they target is minimal.

G-quadruplex aptamers have some common interesting sequence characteristics. The most common base found between the guanosines is thymine, although adenine and cytosine are also observed. Many G-quadruplex aptamers consist of guanines and thymine alone, e.g. d(TGGGg). Thymine is particularly favored as this base has a monocyclic pyrimidine and does not base-pair to guanines. Substitution of thymine with adenine tends to inhibit G-quadruplex structure because adenine is made up of a bicyclic purine that is too bulky for the loop sequence connecting the G-quadruplex. Substitution of thymine with cytosine tends to inhibit G-quadruplex structure due to the competition with surrounding guanosine for Watson and Crick base-pairing.

Additionally, G-quadruplexes may be classified into two topologies - parallel and anti-parallel. Parallel G-quadruplexes contain four strands of nucleic acids in the same direction; whilst anti-parallel G-quadruplexes contain four strands of nucleic acids in opposite directions. The topology of G-quadruplex is usually determined by circular dichroism. Aptamers validated to have G-quadruplex structure usually do not have a defined topology. Actually, the mechanism by which a G-rich sequence folds into parallel or anti-parallel topology remains elusive, and so far no general rule has been reported to accurately predict the topology of a particular G-rich sequence. The folding of G-quadruplex can be controlled by a variety of factors, including but not limited to metal ions, loop sequence, and number of consecutive G-quartets. Since there are many possible structural arrangements of G-quadruplex structures (parallel, anti-parallel, unimolecular, bimolecular and tetramolecular), G-quadruplex structure, originally conceived as a simple rigid motif, actually exhibits remarkable diversity [11]. Such molecular diversity of G-quadruplex aptamers with a similar scaffold can explain the high specificities and binding affinities towards their targets.

G-QUADRUPLEX APARTMER STRUCTURE AND STABILITY

The structure and stability of G-quadruplex oligonucleotides in a given environment was first investigated by studying short sequences derived from telomeric repeats and motifs [36]. Interest in naturally derived G-quadruplexes has since grown considerably, which is not surprising given the plethora of G-quadruplex forming sequences in telomeres, promoter regions, RNA transcripts, and immunoglobulin switch regions [37-39]. Since the discovery of G-quadruplexes in SELEX-derived sequences, similar techniques and approaches have been applied to understand G-quadruplex formation in DNA and RNA aptamers [7]. Examples of G-quadruplex forming aptamers and biophysical characterization of G-quadruplex aptamers have been reviewed [11, 37, 40]. In recent years, important research has focused on understanding G-quadruplex stability and structure through backbone, sugar, and nucleobase substitutions and modifications in the aptamer sequence, yet have not been summarized formally.

The best characterized aptamer in the sense of structure and stability is a 15nt thombin binding aptamer (TBA) which targets human alpha-thrombin [15]. TBA was SELEX enriched to the consensus sequence d(GGTTGGTGGTGGTGG) which was found to inhibit thrombin activity [15]. Thrombin, the serine protease which converts fibrinogen into clottable fibrin in the hemostasis process, importantly contains an anion binding recognition exosite I (fibrinogen binding) and exosite II (heparin binding) [7, 41-43]. It is these two sites that are most vital to the complex with TBA. To illustrate one of these sites a distant view of TBA interaction with exosite I based on X-ray structure is shown in Fig. (2a), with a closeup of TBA in Fig. (2b). Although NMR data does differ from X-ray crystal structure data in observing how the loops interact with the mentioned thrhombin domains, it is well established that TBA interacts with two thrombin molecules while inhibiting only one [7, 41-43]. In its interaction with exosite I, three thrombin amino acid residues form ion pairs with TBA as T7 and T9 associate with a hydrophobic cluster as shown in Figs. (2c and 2d).

The debate over TBA’s molecularity in solution has been a central theme since the beginning of its characterization [5-7, 42, 44-52]. The general consensus was derived from structural work in the 1990s which observed TBA as an anti-parallel unimolecular quadruplex with a chair-like intramolecular structure under physiological conditions [7, 42]. However, a recent study provided evidence that TBA can be bimolecular and parallel in various salt conditions [44]. Circular Dichroism (CD) showed this phenomenon was dependent on aptamer concentration [44]. In addition, lengthening of loop regions and/or the S’ end resulted in strengthening the preference of the modified aptamer for the intramolecular state as compared to the parent TBA [44]. This work illustrated that the physical state of TBA depends on several variables including binding to thrombin and stabilizing cations [69]. When destabilized, TBA is suggested to denature first by breaking G-G base pairing followed by opening of the TGT then TT loops in an exothermic process [51, 65]. Hence, insight into the mechanisms of folding and binding has been gained by atomic substitutions in the TBA backbone, sugars or bases. Because of its promise in clinical applications, TBA has been a focus of attention to better understand its stability and improve its usefulness.

Two key variables relating to structure and stability of TBA, as well as other G-quadruplexes, are the anti or syn glycolysyl torsion angles of the bases relative to the sugars, as well as the C2'- or C3'-endo sugar pucker conformations in the case of RNA Fig. (3). It is well established that deoxyriboguanosines generally favor the syn conformation while riboguanosines favor the anti, but this is not absolute considering that other factors such as sugar pucker influence the conformation [66]. It follows that DNA oligonucleotides favor anti-parallel topology, whereas RNA oligonucleotides favor parallel [66]. For DNA, the anti syn conformation is known as glycosidic bond angles (GBA), can have 16 different positions, and there are 26 permissible looping conformations [70]. The rules governing the relationship between the GBA, loops, strand polarity, and groove were excellently summarized in 2007 [70]. In the anti-parallel strands of TBA, four of the guanines have syn conformation while four have anti in an alternating pattern [5, 6, 46, 71]. In a recent illustration of the importance of the syn or anti conformations of the G-quadruplex, substitution of several dG residues with rG caused conversion back and forth from anti-parallel to parallel
quartets Fig. (4f) [66]. In these experiments, circular dichroism and PAGE revealed the aptamers were bimolecular and parallel, likely to be a single quadruplex composed of two strands [66]. Overall, this work clearly showed the importance of nucleoside conformational preferences to the formation of G-quadruplex and consequences on topology. Also recently, by employing methanocarba-nucleosides, anti or syn conformations were biased so as to understand the importance of the conformations to the stability of TBA [65]. It was found that changing the conformation at G15 especially destabilized the G-quadruplex aptamer and changes at G14 greatly disrupted the structure at large [65]. This study validated the use of locked nucleic acids to deduce the importance of sugar pucker and nucleobase orientation in a given G-quadruplex structure. Another approach, stabilization of TBA was attempted by introducing 2'-deoxy-2'-fluoro-D-arabinonucleic acids into bases to deliberately confer either the syn or anti conformations Fig. (4g) [67]. 2F-ANA conformationally biases the sugar pucker to C2'-endo thus favoring the anti conformation. Using UV thermal melting and CD studies, it was observed that replacement with anti-dGs into TBA stabilized the aptamer whereas insertion of 2F-ANA at the location of syn-dGs destabilized the aptamer [67].

Some recent studies have focused on conformational flexibility. In one example, TBA was substituted with single unlocked nucleic acid monomers at every possible position in the 15 residue chain, thus providing the molecule with more plasticity Fig. (4e) [41]. UV melting analysis and kinetic studies revealed that substitution at various residues made folding more energetically favorable, while substitution of others were unfavorable and diminished thrombin binding [41]. Blood clotting ability also either decreased, or in one substitution, increased. These substitutions concluded mostly what was expected: introducing more flexibility in the quadruplex forming residues and loop regions generally decreased stability and binding, but insertion of a modification at certain locations can have some unexpected advantages. Another study introduced the more rigid modification of thiophosphoryl internucleotide bonds in the chain backbone, which were placed at loop regions of TBA to decrease flexibility and endowed TBA with increased thermostability Fig. (4c) [60]. As expected, stability decreased when substitutions were made in the planar regions. Comparing this with the decrease in thermo stability from unlocked nucleic acids in the loop region confirms the importance of loop rigidity to the overall stability of the TBA molecule [41]. Another example that employed locked nucleic acids showed by NMR, CD, pro-thrombin assays, and structural calculations, that the single LNA modifications changed only the sugar pucker but not the anti/syn conformations which ultimately stabilized TBA Fig. (4d) [61].
The nucleobases are directly responsible for the interactions that comprise the G-quadruplex, and hence full nucleobase substitutions have also been investigated. In one alternative approach to aptamer enrichment (SELEX), a randomly substituted isoguanine (iG) library of the TBA sequence was screened against human alpha thrombin and binding species were sequenced and synthesized Fig. (4a) [55]. When incorporated into particular planar sequences, isoguanine can form a pentamer association around a monovalent cation, whereas guanines can only form tetramers in the quadruplex. Two iG substituted aptamers were found to have substantially higher binding affinities than the parent as determined using Isothermal Titration Calorimetry (ITC), albeit not due to pentamer forming sequences [55]. In contrast, substitution of only one of the Gs in the quadruplex with 6-thioguanine prohibited quadruplex formation entirely Fig. (4h) [68]. This was more dramatic than expected, as in some cases quadruplexes can still form with inosine substitutions [72] In a more comprehensive study, systematic changes were made to the backbone, bases, and sugars Fig. (4) [58]. It was found that changing T’s to U’s had a stabilizing effect, loss of negative charge in the backbone had a destabilizing effect, and hydrogen bond capacity at the 2’ position of the ribose was shown to be of particular importance [58]. In an exclusively computational approach, structure activity relationships of a photo-labile caged TBA aptamer were performed assuming the analogue at the 6th guanosine position Fig. (4a) [53]. Simulations revealed distortions due to the caged TBA which manifested in the loop regions, Hoogsteen binding of guanosines, central monovalent cation positioning, and base stacking. Upon removal of the photo-labile group, a decrease in groove length and G10 in particular contributed the most to TBA formation [53].
In one example which involved an alternative to replacing specific bases in some way, a 5'-5' inversion was introduced to TBA, thus named mTBA d(3'GGT 5'-GGTGGTGGTGG) 3' Fig. (4b) [73]. NMR and CD studies revealed this modification to have three strands in the same direction and one strand in the opposite direction, which may be the reason for its increased thermal stability [73]. Human plasma assays showed a reduced biological activity however. Two years later the same group used Differential Scanning Calorimetry (DSC) and ITC to better understand the inverted TBAs (mTBA) interaction with thrombin [59]. The greater stability of mTBA was attributed to its greater enthalpy value, which may suggest additional hydrogen bonding. The 1.2 stoichiometry of aptamer to thrombin was also confirmed with the ITC experiments [59]. Recently, the crystal structure of mTBA with thrombin was solved [74]. By virtue of a higher resolution (2.15-Å vs. 2.9-Å), the model for mTBA binding to thrombin was clarified. This structural study revealed that both hydrophilic and hydrophobic interactions contribute to binding, including interactions mediated by leucine, isoleucine, and glutamic acid. More significant hydrogen bonding is involved between the TT loops of mTBA binding exosite I as compared to TBA, helping to explain its higher affinity.

The two lateral TT loops and central TGT loops of TBA have also been the target of modifications specifically in order to put their role into context. To this end, loop regions were substituted for other nucleotides, extra nucleotides were inserted and some were truncated in order to study the thermodynamic stability using UV melting curves and DSC [48]. The thermodynamic profiles indicated that different loop sequences and compositions significantly affects stability and the TGT loop in particular. Later, both loops were individually substituted with acyclic thymines, which shortens the distance between the base and phosphate, introduces more hydrophobicity, and creates a chiral center [64]. Structural information was compared to biological activity using NMR, UV, CD, Protein X-ray crystallography (X-ray), Thrombin Time (PT), and Fibrinogen Clot Assays. A correlation was seen between structure and activity as they pertain to substitutions in the TGT loop [64]. Further, it was concluded that substitutions at T3 and T12 have little effect whereas substitutions at T4 and T13 destabilize which was presumed to be due to the loss of a hydrogen bond between the two residues. T3 or T4 substitutions were also shown to have less biological activity than T12 or T13 substitutions, implicating the importance of the 5' TT loop in the inhibition mechanism [64]. Alternative nucleotides, extra thymines and 2-aminopurine were substituted into the loops [57]. DSC and CD techniques were used and thermodynamic parameters for the loops were subtracted from that of one G-quartet stack. The loops were discovered to contribute to the stability due to favorable heat contributions and unfavorable entropy, with all loop changes lowering overall stability [57]. Most recently, the thermal stability of TBA was measured in the context of thrombin binding while exchanging loop thymines with adenines [56]. It was concluded that the loop regions which hydrogen bond to thrombin and bind the intramolecular cation are likely to indirectly stabilize the G-quadruplex.

Finally, a few recent papers have investigated the consequences of TBA substitutions on biological activity and physiological stability. A near 3-fold decrease in IC50 was seen as compared to TBA using thrombin clotting assay when various thymines at the 3,7,9, and 13 positions were substituted with 4-thio-deoxyuridylate Fig. (4a) [54]. A similar increase in the modified aptamers' inhibition in plasma and an increase in platelet activation was observed as well [54]. Structural and thermal stability information would be helpful to better understand the mechanism behind the inhibition as it seems the only point of reference available is previously mentioned work in which single thio-substituted dGs strongly destabilized TBA [68]. Moreover, in order to compare TBA thermal stability and topology with a clotting activity assay, locked nucleic acids were used at G2, T4, G5, T7, and T8 positions Fig. (4d) [62]. The biological activity and stability did not appear to correlate, although conditions used in stability studies are often dramatically different from those in activity assays. The TGT loop, rather than overall stability, was implicated in effecting biological activity most significantly [62]. As for nuclease resistance with regard to TBA, one recent paper attempted capping the 3' end with various 2' and 4' bridged nucleotides Fig. (4d) [63]. These sugar modifications showed a several-fold increase in nuclease resistance in both phosphodiesterase assay and human serum while maintaining binding ability as shown in capillary electrophoresis [63].

Overall, various recent studies have elucidated the consequences of changes made to the various structural regions of TBA which may pertain to similar structural motifs also. Understanding the states of the antidyne or sugar pucker conformations has proven to be useful in explaining causes of the potential loss of hydrogen or ionic bonding. Through the replacement of dG with dG residues [66], the employment of methanocarba-nucleosides [65], and the use of 2'-F-ANA [67], the rules of quadruplex formation as proposed theoretically by Webb da Silva were understood using the real example of TBA [70]. The introduction of increased flexibility or rigidity with unlocked [41] or locked nucleic acids [61] has provided insight into the key regions of the sequence that provide the most stability. Base substitutions have also provided a greater understanding of overall stability and formation [53, 58]. Importantly, a superior functional aptamer was created by introducing a 5'-5' inversion [59, 74]. Finally, significant perturbations to the loops have also provided some illustration of their contribution to the overall topology, which can be compared to the proposed rules of G-quadruplex formation - at least for a chair-like, one quadruplex stack DNA aptamer conformation [48, 56, 57, 66, 70]. It is hoped that such a diverse panel of changes be instigated for other parent aptamers so comparisons can be made in the future.

EFFECTS OF CATIONS ON G-QUADRUPLEX APTAMER STRUCTURE AND STABILITY – THE CASE STUDY OF TBA

G-quadruplex topology and stability was first understood from the perspective of telomeric sequences, and from these examples some generalities were drawn such as: 1) cation ability to induce quadruplex formation follows the trend K+>Rb+>Na+>Cs'>Li' with mechanistic differences being the most understood between K+ and Na+. 2) many cations move in and out of the quadruplex core if not obstructed, 3) K+ values for divalent cations are significantly greater than those for monovalent cations and 4) stabilizing abilities of divalent cations relate in the order Sr2+>Ba2+>Ca2+>Cs+>Mg2+ [40]. A number of studies have investigated the influence of cations on DNA aptamer structure, and TBA is again the most deeply studied model for this aspect of aptamers that can act as a case study. Not only do cations have an important role in stabilizing quadruplex aptamers, but this property has even been exploited for the purpose of cation detection. The TBA aptamer itself has also been used to detect Hg2+ and Pb2+ [75]. As a further example, aptamers have been generated for the detection of K+ using fluorometric and nanosilver resonance scattering methods [76, 77]. To some degree, these detection methods owe their approach to previous work in understanding the relationship between cation and DNA G-quadruplex aptamer.

G-quadruplexes owe their remarkable stability to a conglomeration of forces including hydrogen bonding, Van der Waals, stacking, hydrophobic, and ionic interactions. More unique, however, is the stability cations lend themselves to quadruplex structure, which is thought to be due to an association with O6 carbonyl groups between the planes [78, 79]. In the case of aptamers however, and for TBA in particular, the placement of the stabilizing cation is still debated [49, 80]. Also central to the discussion of TBA is the dependence on, and stoichiometry of, different cations which are of special interest for those cations available under physiological conditions...
dictions. Some of the initial structural work on TBA asserted that TBA depends on cations such as potassium for its chair-like active fold [5, 6, 46, 81-83]. TBA was assessed using NMR in the presence of Na" and K" [81]. In this study 100 mM Na" alone showed non-exchangeable protons, thus TBA lacked the chair conformation whereas K" profoundly changed the conformation to the chair fold. This was shown to occur in a sequential 1:2 TBA to K" stoichiometry with greatly differing binding strengths, with the first binding site being between the TT loops Fig. (5b) and the second adjacent to the TGT loop Fig. (5e) [81]. Afterwards, it was discovered that Pb" is more potent in inducing the TBA fold and showed a 1:1 stoichiometry as compared to K" [84]. The same group later provided evidence that the Pb" binding site was between the two quartets Figs (5b) [85]. Further studies of the structure of the potassium TBA complex by NMR confirmed the necessity for K" in the chair type folding of TBA as had been claimed in the past [5, 6, 46, 81]. It was also found that the first binding K" is sufficient for chair fold formation and the second K" that binds with increased concentration influences the TGT loop specifically.

More recently, a comparison between NH4, K" and Na" ions in terms of their localization within TBA was evaluated by NMR and other spectroscopic methods and conflicting conclusions were made about topological and stoichiometric aspects [86]. The topology formed in the presence of NH4 was similar to K", Sr", and Pb" [86]. More importantly, binding between the two G-quadruplexes with the carbonyl oxygens was argued to be the location for these ions rather than between TT loops or adjacent to the TGT loop Fig. (5). Out of plane bending of G2 and G11 was seen, which supports the argument for inter-quartet localization. A unimolecular fold at a wide range of K" concentrations was also seen which contrasts with Marathias and Bolton [49]. Also importantly, a unimolecular fold was seen at a wide range of K" concentrations in contrast to two previously mentioned reports [44, 64]. It was also concluded that Na" folds and unfolds most rapidly and K" has a slow unfolding reaction, which implies an ephemeral interaction with Na". In another example in which several different cations were examined, the effects on the conformation, unfolding, complex formation, and molar volume were assessed with CD, DSC, ITC, density and acoustical techniques [87]. K", Ba", NH4, Sr", and Ba" were found to form stable intramolecular complexes above 25°C, while Li", Na", Cs", Mg", and Ca" were found to form weak complexes at low temperatures. The results were noted to coincide with the atomic radii of the cations.

A mass spectrometric approach, Electrospray Ionization (ESI), has also proved a viable way to detect the formation of G-quadruplex in gas phase along with bound cations [90]. Asking if metal ion binding is specific, adenine substituted TBA was used as a control. Li", Na", K", and Cs" ions were found to bind to the controls, however, enhanced adduct formation for K" as opposed to the control implied specific binding. The ions which showed enhanced binding compared to control – Sr", Pb", Ba", and K" - agree with the Kankia et al. correlation of ionic radii to G-quadruplex binding [87]. In contrast, a 2010 report employed ESI, sequencing with infrared multiphoton dissociation (IRMPD), and CD to a TBA study investigating binding modes of K", Na", Pb", and Cs" [80]. IRMPD is able to sequence TBA by the breaking of the phosphate backbone in order to deduce ion association based on the fragments. This work concluded that K" is first positioned in between the two TT loops and the second (1:2) K" positions adjacent to the TGT loop as proposed by Marathias et al. Fig. (5) [49]. The significantly higher binding strength of K" as compared to other cations was also confirmed here in accordance with past work [5, 6, 46, 81, 82].

![Fig. (5). Tentative locations for cations in DNA aptamer G-quadruplexes. a: interplanar, b: subplanar, c: supersubplanar, d: intraplanar.](image)

In order to determine TBA’s ability to fold into an anti-parallel intramolecular structure in the absence of a stabilizing cation, CD studies were performed in conditions of low temperature, molecular crowding, and thrombin as opposed to K" [69]. In this study, Na" was found to induce quadruplex formation at 4°C but hardly induce formation at 25°C. Quartet formation was induced by molecular crowding with PEG 200 as well as with thrombin alone. It was proposed however, that TBA has low inhibitory activity without K" [69, 91]. Finally, in a study which defined some of the factors that favor the G-quadruplex fold, different cosolvents and solutes were added to TBA environment [92]. This article showed that solutes which decrease DNA duplex stability have an opposite effect for
quadruplex formation, and the dielectric constants of the solvents have a large impact on stability, which would be consistent with cation-induced stability.

G-QUADRUPLEX DNA APTAMER THERAPEUTIC APPLICATIONS

The majority of aptamers developed for therapeutics are not necessarily G-quadruplex. The reader interested in therapeutic aptamers generally including the most successful of the class Pegaptanib/Macugen used clinically is pointed to a recent review [93]. In the section below we focus on therapeutic DNA G-quadruplex aptamers specifically.

G-quadruplex DNA Aptamers Against Infective Agents

A major class of therapeutic G-quadruplex DNA aptamers against infective agents are those against HIV. In an early study, a library of all possible octanucleotide phosphorothioate sequences was screened and d(TTGGGGTT) was found as an inhibitor of HIV infection that acts by binding to envelope protein Gp120 thereby inhibiting cell-to-cell and virus-to-cell infection [22]. Size-exclusion chromatography revealed that the oligonucleotide was tetrameric, and circular dichroism suggested that the phosphorothioate adopted a parallel G-quadruplex structure [22]. The aptamer was shown to have synergistic anti-HIV effects together with AZT against HIV in human cell lines [94]. Various hexanucleotides, based around the d(TGGGAG) Hotoda sequence also exhibited strong anti-HIV properties [23, 95, 96]. A particularly exciting extension of these ideas are to use anti-gp120 aptamers (although in this case 2F-substituted RNA aptamers) to deliver siRNA into HIV infected cells [97].

Subsequently, an aptamer consisting only of deoxyguanosine and thymidine was shown to strongly inhibit HIV integrase with an IC50 in the nanomolar range [98]. Sequence change derivatives revealed that loop-ion interactions were critical for stability and inhibitory activities against the HIV integrase, and also resulting in an aptamer more stable than the closely related thrombin binding aptamer [99]. Further structure-activity studies using NMR and molecular modeling showed that the K+-bound form as an antiparallel intramolecular fold with a pair of G-quadruplexes was critical to inhibitory activities [100]. DNA aptamers which were derived from aptamers against HIV-1 RNase H domain of reverse transcriptase inhibitors, which share a similar topological organization, were also shown to inhibit the HIV integrase [101]. These aptamers were later shown to adopt a stable dimeric G-quadruplex structure in the presence of potassium ions [24].

The HIV reverse transcriptase has also been targeted with DNA aptamers by a number of independent groups. Although the first SELEX experiments of DNA against reverse transcriptase did identify G-rich aptamers [102], it was not proven at that time whether they were indeed G-quadruplexes. More recent studies did identify related aptamers to be G-quadruplex [26]. DNA thioaptamers selected against reverse transcriptase were identified with a Kd of 70 nM which also inhibited HIV replication in cell culture [103]. In yet another selection, DNA aptamers were found with a distinct run of consecutive G-rich sequences suggesting these may also have G-quadruplex structure [104].

A study selecting aptamers against the SARS coronavirus helicase showed that both G-quadruplex and non-G-quadruplex aptamers were identified in the same selection procedure [14]. The non-G-quadruplex did inhibit helicase activities, whereas the G-quadruplex aptamers did not, implying a clear structure-function relationship between the G-quadruplex nature of the aptamer and its inhibitory activities.

Aptamers have been selected against polypephosphate kinase 2 from Mycobacterium tuberculosis and both G-quadruplex and non-G-quadruplex DNA aptamers were identified [13]. Mutational analysis revealed that the G-quadruplex motif was essential for inhibition. Aptamers against MPT64 were found to have several consecutive G-rich sequences suggesting these may also have G-quadruplex structure [109].

G-quadruplex DNA Aptamers to Control Coagulation

We have discussed the thrombin-binding aptamers in much detail in terms of dissecting the structure of the thrombin-aptamer complex in the sections above. A recent review has recapitulated progress with regards to clinical trials for nucleic acid aptamers generally as antithrombotic agents [110], incorporating aptamers against a variety of targets that are not necessarily G-quadruplex and hence beyond the scope of this review. The thrombin-binding aptamer itself was halted after Phase I clinical trials, but the many variations such as all the chemical modifications discussed prior and/or bivalency [111] may prove more effective.

G-quadruplex DNA Aptamers for Cancer

AS1411 is a G-quadruplex DNA aptamer that is undergoing Phase II clinical trials as an anticancer agent, for renal cell carcinoma and myeloid leukemia. The discovery of AS1411 was to an extent coincidental – a result of using the “purine motif” approach of using triplex-forming oligodeoxynucleotides to mediate gene expression [112]. This led to sequences consisting entirely of guanine and thymine, one of which was a 26-mer that became AGRO100 while being developed at Aptamera, then was renamed as AS1411 after acquisition by Antisoma. In May 2011, AS1411 was acquired by Advanced Cancer Therapeutics (ACT), and renamed ACT-GRO-777. AS1411 binds to both nucleolin and to nuclear factor-kB (NFκB), which is involved in activation of the critical transcription factor NFKB [113]. Phase I clinical trials were completed in 2006 (a total of 30 patients) [112], with Phase II trials against acute myeloid leukemia showing positive interim data. Recent research has extended the use of AS1411 as a carrier for the delivery of cancer drugs specifically into cancer cells by nucleolin mediated internalization [114].

Signal transducer and activator of transcription (Stat3) is an oncogene that is activated in many human cancers. A G-quadruplex aptamer was found to inhibit DNA binding activity of Stat3 with an IC50 of 7 μM [115]. Furthermore, the G-quadruplex aptamer was delivered efficiently to the cytoplasm and nucleus, and was shown to suppress Stat3 oncogenic pathways. These G-quadruplexes were further developed and shown to suppress tumor growth in a non-small cell lung cancer model in mice [18].

G-quadruplex DNA APTAMERS FOR SKELETAL DISEASE

Sclerostin is an extracellular negative regulator of bone growth that is major therapeutic target for osteoporosis. Parallel to a variety of efforts to develop antibodies against sclerostin (which recently completed Phase II clinical trials), DNA aptamers were also selected against sclerostin as an alternative approach and found to have parallel G-quadruplex structure [12]. These aptamers were effective in a cell culture model to inhibit sclerostin’s antagonistic effect on Wnt signaling, and are under further development as an alternative approach to inhibiting sclerostin for skeletal disease.
G-QUADRUPLE DNA APTAMER DIAGNOSTIC, DETECTION AND SENSING APPLICATIONS

The G-quadruplex plays a significant role in mediated analyte detection by aptamers, which are commonly referred to as aptasensors. As recently as 2010 aptasensors were reviewed in the context of label and substrate free systems and nucleic acid nanostructures and machines were reviewed as well [116, 117]. Most recently, a review focused on aptamer based detection with excellent descriptions of the technical assay strategies associated with this approach [118]. However, design and use of aptasensors in the context of the G-quadruplex itself has not been summarized and therefore we integrate several recent papers here.

Important to the topic of aptasensor systems is the use of metallic beads, electrodes, and quantum dots which enable more sensitivity over colorimetric systems and allow for signal amplification schemes. One important development is the use of a horseradish peroxidase (HRP) mimicking DNAzyme system whereby the iron containing porphyrin hemin binds to an 18 base pair, K⁺ dependent aptamer d(GTGGGTAGGGCGGGTTGG) which was selected against N-methylmesoporphyrin IX [32]. Hemin’s inherent peroxidase activity was then found to significantly increase when in complex to the aptamer, hence HRP mimicking DNAzyme was introduced [32]. It was later realized that the RNA version of PS2.M (rPS2.M) also showed enhanced peroxidase activity, demonstrating that RNA counterpart aptamers can have similar structure and function [119]. The intramolecular guanines were also shown to play an important role in the peroxidase activity [120]. Later, the structural and stability properties of PS2.M were further studied with spectroscopic methods and NMR [121]. This work revealed that Na⁺ and Pb²⁺ induce unimolecular structures but the presence of K⁺ forms both unimolecular and multistranded types. It was concluded that one Pb²⁺ cation confines the three quadruplex fold, and the active complex in the presence of K⁺ may result from a multistranded parallel type structure.

Because of its biological significance, thrombin has been the analyte in many assays, some of which employ the HRP mimicking DNAzyme system. In two such examples which detected thrombin in the nanomolar range, metallic nanoparticles were used in the detection scheme [122, 123]. These strategies linked TBA to PS2.M in some way such that peroxidase activity was affected upon thrombin binding. More complicated schemes based on the initial event of the TBA folding upon thrombin binding have also been used [124]. One interesting approach involves a TBA and hemin complex which also possesses some peroxidase activity [125]. The chair conformation supported by the presence of K⁺ was shown to have a weak interaction with hemin, but in the presence of thrombin the interaction changed so that peroxidase activity was significantly higher, allowing colorimetric detection at 20 nM substrate. In another similar example in which TBA itself binds hemin to enhance peroxidase activity, thiamine reduction was exploited for a fluorescence signal enabling a detection limit in the picomolar range [126].

Alternatively, a few thrombin detection schemes do not employ peroxidase activity but rely on quadruplex formation changing atomic proximities thus leading to a signal. Recently, a 50 nM detection of thrombin using fluorescence quenching also used CD to confirm quadruplex formation in the process [127]. Furthermore, electrochemical detection has recently detected thrombin using gold electrodes [128, 129] and gold nanoshells [130]. In one strategy, thrombin binding and quartet formation leads to dissociation of double stranded DNA thus freeing a quenched fluorescent label [129]. The gold nanoshell approach uses Raman spectroscopy in a label free, direct measurement [130]. Raman spectroscopy senses the sugar pucker conformation, thus further emphasizing its importance to the G-quadruplex. Apart from thrombin, the cancer-associated nucleolin protein and an HIV-1 integrase have also been detected with aptamers AS1411 and T30695, respectively, while making use of their own peroxidase activity [131], AS1411, which binds HeLa cell expressed surface nucleolin, was used for in situ sensing. This work also importantly asserts that AS1411 requires K⁺ for folding while cations actually inhibit T30695 quadruplex formation. Further structural and stability work on these aptamers would offer a needed comparison to the better characterized TBA.

Regarding small molecules, adenosine, AMP, and ATP have also been the analyte in some recent assays. Adenosine is key to kidney function [132] and has been associated with cancer [133]. The initial selection of the adenosine/ATP aptamer used sequence substitutions and nucleotide analogues to demonstrate G-quartet formation [35]. Upon sequence comparison, a two quartet structure similar to TBA was observed, although with longer loop regions which include a hairpin on one end and a duplex on the other [35]. In one example of adenosine detection, an anti-flavin adenosine dinucleotide-dependent glucose dehydrogenase (FADGDH) aptamer was attached to a scaffold containing the adenosine aptamer [134]. Upon binding of adenosine, quartet formation brings FADGDH closer to the electrode to confer a signal. In a similar strategy, quartet formation upon adenosine binding used gold nanorods and Plasmon Resonance Absorption (PRA) was monitored to detect analytes in the low nM range in rat brain tissue [135]. Similarly, adenosine binding aptamer (ApT₃₀₆₉₅) and platelet derived growth factor (PDGF) binding aptamer (ApTʹ₃₀₆₉₅) were used to bind adenosine and boost signal respectively [136]. Here, folded (ApT₃₀₆₉₅) displaces unfolded (ApT₃₀₆₉₅) on the surface of gold nanoparticle oligo labeled oligos. Further, two recent strategies use a PS2.M/hemin HRP mimetic system in conjunction with the adenosine aptamer [137, 138]. In one, a complicated system of “logic gates” allows for the biocatalytic regulation of thrombin activity with inputs such as ATP or cocaine [137]. In a more simple example, a duplex using (ApT₃₀₆₉₅) dissociates upon adenosine binding which causes gold nanoparticle aggregation [139]. Finally, detection of adenosine deaminase activity was achieved by allowing conversion of adenosine to inosine to confer a flexible structure freeing hemin and K⁺ thus forming a separate quadruplex for an electrochemical response [140].

DNA sequences have also been a target for some recent detection strategies. Again, exploiting the fact that quadruplexes form only under certain conditions, hairpin DNA structures in a “caged” form containing the sequence for a particular aptamer are forced into folding upon analyte binding, which is the case in a recent method to detect DNA, AMP, and Hg²⁺ [141]. DNA, AMP and glucose were also detected recently in a similar manner by the same group [142]. In an interesting PCR based approach, “padlock probes” that straddle single nucleotide polymorphisms (SNPs) were ligated, amplified, and introduced to hemin to confer G-quartet in HRP-mimicking DNAzyme thus giving rise to signal [124]. The HRP DNAzyme system was also used in a way in which PS2.M was extended with sequences complementary to a target such that the PS2.M quadruplex would form upon a nucleotide mismatch [143]. HRP DNAzyme was also used in a system using quantum dots that allow for a multiplex detection of different DNA sequences using chemiluminescence resonance, energy transfer (CRET)[144]. Hg²⁺ and ATP were also detected using this system[144]. Some small molecules have also been targeted recently which including hemin itself based on the PS2.M system [145]. Ochratoxin, a mycotoxin found in foods, was detected with ochratoxin aptamer [33] immobilized on gold nanoparticles [146]. Ochratoxin aptamer forms quadruplexes upon ligand binding, which was exploited to aggregate gold nanoparticles for detection[146].

Regarding aptamer-mediated cation detection, K⁺ has been detected using a designed K⁺ binding quartet forming sequence [34] in which nanosilver aggregates upon quartet formation [77] and by displacement of K⁺ from the quadruplex sequence by a complementary strand [76]. This method was also used for DNA sequence detection. Pb²⁺ has also been a target analyte based on the PS2.M
system whereby the stable K⁺ and hemin bound aptamer is disrupted by Pb²⁺, thereby abrogating peroxidase activity [147].

**FUTURE PERSPECTIVES**

It has now been over two decades since the original pioneering nucleic acid aptamer selections. As the original patents for the SELEX process expire, there will likely be a reinvigoration of aptamer research with therapeutic and diagnostic ambitions [148, 149]. As such, selection of aptamers against new targets will no doubt lead to an increasing diversity of DNA G-quadruplex aptamers over coming years.

The structure of the thrombin binding aptamer–thrombin complex has been particularly instructional to elucidate folds, ion binding, grooves and binding mechanisms of G-quadruplex aptamers. However, we critically need high-resolution structures of other G-quadruplex aptamers in complex with their targets to answer the question of whether TBA is typical or atypical, and to help us answer the question of how G-quadruplex aptamers manage such high target specificities and binding affinities. New structures will help G-quadruplex aptamers meet their therapeutic promise and analytical potential, leading to a new chapter in DNA G-quadruplex aptamer research.

**ACKNOWLEDGMENTS**

We thank Wong Ka Wai and Lam Chun Ming for their help in the preparation of the figures in this review. The communicating author’s work related to applications of nucleic acid aptamers is funded by the Hong Kong Research Grants Council (RGC) under GRF grants HKU 777109M and HKU/776108M.

**REFERENCES**

[38] Weisman-Shomer P, Fry M. QUAD, a protein from hepatocyte chromatin that binds selectively to guanine-rich quadruplex DNA. J Biol Chem 1993; 268: 3306-12.


Smirnov IV, Kotch FW, Pickering IJ, Davis JT, Shafer RH. Pb


