The emerging structural complexity of G-quadruplex RNAs

MICHAEL T. BANCO and ADRIAN R. FERRÉ-D'AMARÉ

Biochemistry and Biophysics Center, National Heart, Lung, and Blood Institute, Bethesda, Maryland 20892-8012, USA

ABSTRACT

G-quadruplexes (G4s) are four-stranded nucleic acid structures that arise from the stacking of G-quartets, cyclic arrangements of four guanines engaged in Hoogsteen base-pairing. Until recently, most RNA G4 structures were thought to conform to a sequence pattern in which guanines stacking within the G4 would also be contiguous in sequence (e.g., four successive guanine trinucleotide tracts separated by loop nucleotides). Such a sequence restriction, and the stereochemical constraints inherent to RNA (arising, in particular, from the presence of the 2′-OH), dictate relatively simple RNA G4 structures. Recent crystallographic and solution NMR structure determinations of a number of in vitro selected RNA aptamers have revealed RNA G4 structures of unprecedented complexity. Structures of the *Sc1* aptamer that binds an RGG peptide from the Fragile-X mental retardation protein, various fluorescence turn-on aptamers (Corn, Mango, and Spinach), and the spiegelmer that binds the complement protein C5a, in particular, reveal complexity hitherto unsuspected in RNA G4s, including nucleotides in *syn* conformation, locally inverted strand polarity, and nucleotide quartets that are not all-G. Common to these new structures, the sequences folding into G4s do not conform to the requirement that guanine stacks arise from consecutive (contiguous in sequence) nucleotides. This review highlights how emancipation from this constraint drastically expands the structural possibilities of RNA G-quadruplexes.

Keywords: G-quartet; G-tetrad; NMR; tetraplex; X-ray crystallography

INTRODUCTION

In 1910, the Swedish biochemist Ivar Bang reported that high concentrations of guanylic acid form a gelatinous material in aqueous solution. Half a century later, Gellert et al. (1962) deduced from fiber diffraction experiments on this material that four guanines could arrange into a cyclic, Hoogsteen-paired arrangement, the canonical G-quartet (or G-tetrad, Fig. 1A), and that vertical stacks of such G-quartets could result in a four-stranded structure termed the G-quadruplex (G-tetraplex, G4). Formally, at least two stacked G-quartets are needed to form a G-quadruplex. G4s appear to be widespread in biology (Agarwala et al. 2015; Fay et al. 2017; Saranathan and Vivekanandan 2019; Varshney et al. 2020). Analysis of the human genome revealed over 700,000 putative G4-forming sequences (Chambers et al. 2015). Additionally, recent transcriptome-wide studies identified numerous putative G4-forming sequences (Kwok et al. 2016; Murat et al. 2018; Yang et al. 2018; Sauer et al. 2019; Lee et al. 2020), some of which have been validated biochemically (Kumari et al. 2007; Martadinata and Phan 2009; Arora and Suess 2011).

At present, there are 246 G4 entries in the structural database (PDB, curated manually), of which 229 are DNA (or DNA–protein complexes), and the remaining RNA (or RNA–protein complexes). DNA and RNA G4s both exhibit polymorphism (that is, one sequence may form several different G4-arrangements) and share general structural principles. G4s can be formed by either a single nucleic acid chain (unimolecular) or multiple nucleic acid chains (bimolecular, trimolecular, tetramolecular, etc.) The placement of the four guanine carbonyl oxygens near the fourfold axis of the G-quartet generates an electronegative central pore, which is stabilized by coordination of a cation. This cation is most commonly K^+, but NMR and crystal structures exist where the axial cation is Na^+, Cs^+, Sr^{2+}, Ba^{2+}, or even NH_4^+. Cation coordination by G-quadruplexes was reviewed by Bhattcharyya et al. (2016). The guanines forming the G4 structure can be in the *anti* or the *syn* conformation (Fig. 1A), with the latter only possible if steric clash between the ribose and the base is relieved by sugar puckering. The pucker itself is either (predominantly) C2′-*endo* or the C3′-*endo* (Fig. 1B). In the case of the DNA double helix, these correspond to B-form and A-form, respectively, and reflect differences in hydration. For the RNA duplex, the steric block of the ribose 2′-OH imposes the A-form (C3′-*endo*) pucker (Saenger 1984). However,
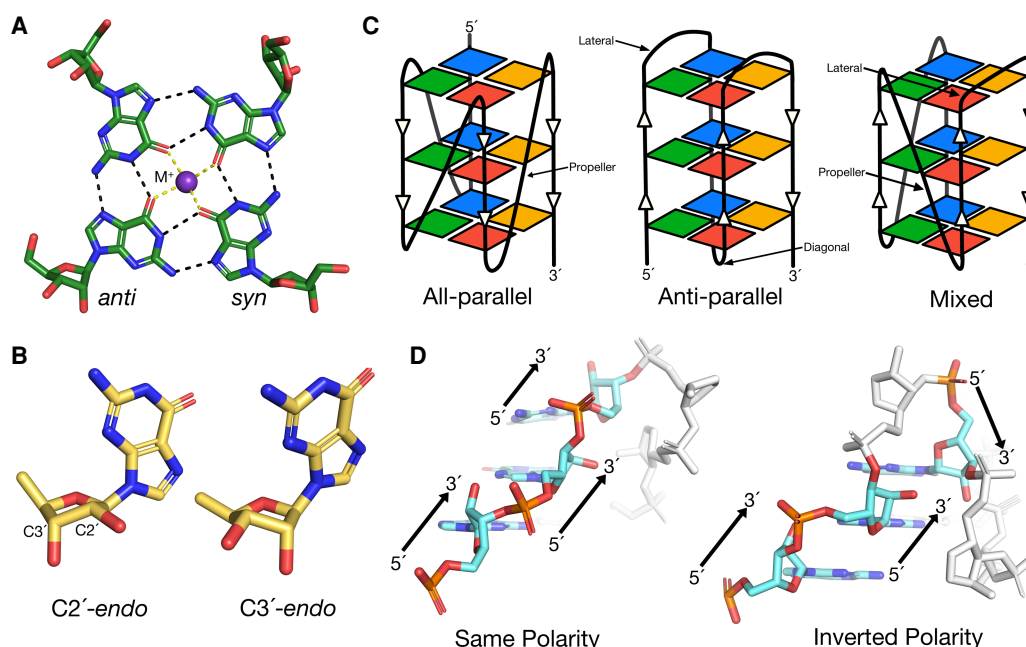

FIGURE 1. General structural characteristics of G4s. (*A*) The cyclic arrangement of guanines in a canonical G-quartet or G-tetrad. G-quadruplexes (G4s) arise from stacking of two or more G-quartets. The glycosidic bond conformations *anti* and *syn*, as well as a cation occupying the central pore are noted. Black and yellow dashed lines denote hydrogen bond and metal ion coordination, respectively. (*B*) As in other nucleic acid structures, the pentose pucker can be C2′-*endo* or C3′-*endo*. In these two examples, the glycosidic bond angles are *syn*. (*C*) Various connectivities of G4 structures, which include all-parallel, anti-parallel, and mixed. Diamonds represent guanine bases. Four colors (red, orange, blue, green) denote the stacked guanines of three successive G-tracts or stacks. The loops connecting the all-parallel guanine stacks are of the propeller type. Lateral and diagonal loops are noted in the anti-parallel G4. Outline arrowheads in the backbone denote 5′-to-3′ chain direction. (*D*) Different local strand polarities (directions) of guanines in G-tracts. Three guanines of a G-tract are highlighted in cyan. In the *left* panel, the 5′ to 3′ direction of the riboses of the three guanines is the same. In the *right* panel, the direction of one of the guanines is opposite.

for nonhelical conformations, the most important correlate of the pucker is the distance between the two flanking phosphates, which is 7.0 Å and 5.9 Å for C2′-*endo* or the C3′-*endo* puckered riboses, respectively (Murray et al. 2003).

The sequential arrangement of their guanine tracts will determine the connectivity of the G4 structures, which are categorized as parallel, anti-parallel, or mixed (Fig. 1C). This refers to the relative 5′-to-3′ direction of the four strands that make up the quadruplex. Depending on the connectivity (not "topology"; no knotted G4s have been described; all known G4 have the topology of a simple line, or multiple simple lines) of the G4 structure, the guanines can stack on each other with the same or inverted 5′-to-3′ polarity (Fig. 1D). Stacking of nucleotides with inverted polarity is commonly observed in nonquadruplex RNA, for instance, when cross-strand stacking occurs within a helix (Jhunjhunwala et al. 2020).

Nucleotides that are not part of a quartet and connect successive G-tracts (arrangements of stacked guanines parallel to the fourfold axis of the G4), form the loops of the G4 structure. There are three common types of loops: propeller, lateral, and diagonal. Propeller loops connect adjacent parallel G-tracts. Lateral and diagonal loops con-

nect adjacent and nonadjacent antiparallel G-tracts, respectively (Fig. 1C). Mixed connectivity G4s can contain a combination of loop types.

Two simple G4 nucleic acids illustrate how the interplay of structural principles described above gives rise to stable G4 arrangements. First is the unimolecular DNA G4 structure formed in the nuclease hypersensitivity element III$_1$ region of the c-*Myc* oncogene (Ambrus et al. 2005). Overexpression of c-Myc is associated with various cancers (Dang 2012). The *c-Myc* G4 structure functions as a repressor of the gene (Siddiqui-Jain et al. 2002; Hurley et al. 2006), which has made it an attractive alternative target for new anti-cancer therapeutics (Hurley et al. 2006; Calabrese et al. 2018; Carvalho et al. 2020). This G4 is an all-parallel quadruplex with three canonical G-quartet tiers (Fig. 2A). All the guanine glycosidic bonds are *anti*, and most of the sugars are in the C2′-*endo* conformation. All loops connecting the guanine stacks are of the propeller type and consist of one or two unpaired nucleotides. As in the vast majority of G4s (Chung et al. 2015; Winnerdy et al. 2019), the loop connections are all right-handed. Second is the G4 structure formed by the telomeric repeat-containing RNA (TERRA) sequence. Although mammalian telomeres were thought to be transcriptionally
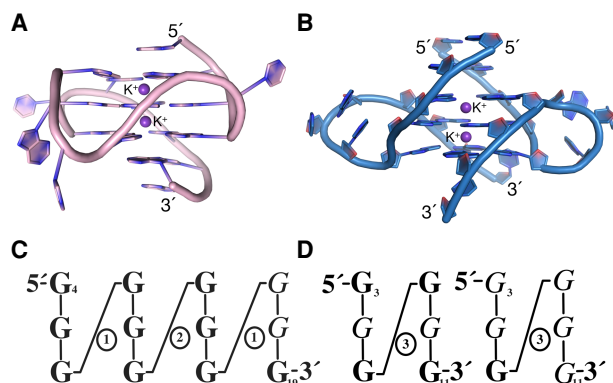
FIGURE 2. Simple G4 structures of *c-Myc* and TERRA. (*A*) Three-tiered, unimolecular G4 structure adopted by a sequence from the *c-Myc* promoter (PDB:1XAV; Ambrus et al. 2005). (*B*) Three-tiered bimolecular RNA G4 structure of TERRA (PDB:3IBK; Collie et al. 2011). (*C*) Schematic of the *c-Myc* G4, highlighting canonical DNA G4 structural features. The guanines of the G4 are *anti*, predominantly C2′-*endo*, and are in an all-parallel connectivity, with mono- and di-nucleotide propeller loops. See Table 1 for code. In all figures, residue numbers correspond to those in the PDB depositions. (*D*) Schematic of TERRA illustrating canonical RNA G4 structural features. The guanines adopt the *anti* conformation are connected in a parallel connectivity and are predominantly in the C3′-*endo* sugar pucker.

silent, long transcripts that include the repeat sequence r(UUAGGG) were identified more recently (Azzalin et al. 2007). Solution NMR structure determination of a stable species formed by TERRA in the presence of K$^+$ revealed a bimolecular G4 with three G-quartet tiers (Fig. 2B; Martadinata and Phan 2009; Collie et al. 2010). The symmetric structure contains two trinucleotide propeller loops, all glycosidic bonds are *anti*, and most riboses adopt the C3′-*endo* conformation.

In the past decade, crystallographic structure determination of RNA aptamers evolved in vitro to recognize various ligands revealed that some of them possess G4s of unprecedented complexity. These structures expand the known range of structural possibilities of RNA G4s and are the focus of this review. A number of solution and crystal structures of DNA G4s have been solved containing noncanonical structural features, such as bulges, vacancies, inclusions from loop nucleotides, and non-G tetrads (Lightfoot et al. 2019). All of these G4 features have now been observed in RNA aptamers and are discussed below. G4 structures adopted by several RNA aptamers challenge the received wisdom regarding the structural scope of RNA G4s. The sequence context of these RNA G4s often does not conform to the canonical sequence motif ($G^{\geq 2}-N^X-G^{\geq 2}-N^X-G^{\geq 2}-N^X-G^{\geq 2}$). Recent reviews provide an overview of the basic structural principles of RNA G4s (Malgowska et al. 2016), and the structural diversity of G4s in general (Lightfoot et al. 2019). Here, we emphasize insights into the potentially much higher complexity of RNA G4s from recently reported structures of in vitro select-

ed aptamers. Other aspects of G4 structure, chemistry and biology, including recognition of G4 by proteins and small molecules, have been extensively reviewed elsewhere (by, among others, Sissi et al. 2011; Mendoza et al. 2016; Hänsel-Hertsch et al. 2017; McRae et al. 2017; Neidle 2017; Sauer and Paeschke 2017; Varshney et al. 2020).

A GRAPHICAL SHORTHAND FOR G-QUADRUPLEXES

Typically, in schematic representations, G4s are drawn as stacks of four squares or diamonds (G-quartet, each square depicting a guanine) connected by lines (loops, Fig. 1C). Although this suffices to indicate the number of quartets and their connectivity, additional stereochemical information is not immediately apparent. The Leontis and Westhof (2001) symbols are widely used to represent noncanonical base pairs, but are not informative for G-quadruplexes, as G-quartets have but one (cyclically iterated) base-pairing scheme: Hoogsteen to Watson–Crick. We propose a graphical shorthand that captures salient structural features of G-quadruplex nucleic acids (Table 1). Structural similarities and differences between the *c-Myc* and TERRA G4s, in particular, how variation of sugar pucker concentrates at the ends of G-stacks, are readily apparent using this representation (Fig. 2C,D).

In our convention, each row represents the nucleotides of a quartet, and columns indicate nucleotide stacks. Mixed-sequence and noncanonical quartets are included in the schematics only when these quartets are stacking directly on a G-quartet. The 5′ and 3′ boundaries of the schematics are the first and last nucleotide that are part of a quartet (canonical or mixed). Strand polarity inversions are indicated as upside-down letters for quartet-forming nucleotides with respect to the 5′-most nucleotide in the schematic. Lines connecting the quadruplex nucleotides are loops (connections between stacks) and bulges (interruptions within stacks). The number of nucleotides in loops and bulges are indicated within a circle. Expansions of a quartet into a pentad, hexad, etc., by inclusion of a loop

TABLE 1. Legend for symbols in G4 schematics

Symbol	Descriptor
Upper case nucleotides (G/A/U/C/T)	*Anti* conformation
Lower case nucleotides (g/a/u/c/t)	*Syn* conformation
Bold nucleotides (**G/A/U/C/T**)	2′-*endo* sugar pucker
Italicized nucleotides (*G/A/U/C/T*)	3′-*endo* sugar pucker
Inverted nucleotides	Inverted strand polarity
⊕	Loop lengths
<	Bulged nucleotides

nucleotide in the plane of the quartet are denoted by drawing a loop nucleotide adjacent to the quartet row.

STEREOCHEMICAL CONSTRAINTS ON SIMPLE RNA G-QUADRUPLEXES

Until the NMR and cocrystal structures of the Fragile-X aptamer (Phan et al. 2011), it had been received wisdom that RNA G4s conform to three general stereochemical rules. First, all guanines involved in G-tetrads adopt the *anti* conformation. Second, RNA G4 structures are limited to a parallel connectivity. Third, the ribose moiety in RNA G4s preferentially adopts the C3′-*endo* sugar pucker. Although C2′-*endo* ribose puckers have been observed in RNA G4s, the C3′-*endo* pseudorotamer was thought to be preferred. For instance, the structure of TERRA contains a mixture of C2′- and C3′-*endo* puckers with the majority of riboses exhibiting the latter (Fig. 2D). The sugar puckers were proposed to result from steric clash between the purine bases and 2′-OHs of guanines forming the G-quartets (Martadinata and Phan 2009). It has been noted that the 2′-OH, as well as other features of RNA, contribute to the structural stability of RNA G4s, resulting in melting temperatures substantially higher than those of their DNA counterparts (Arora and Maiti 2009; Joachimi et al. 2009; Agarwal et al. 2012).

THE FRAGILE-X APTAMER

The fragile-X mental retardation protein (FMRP) is an essential regulatory RNA-binding protein that is associated with several human disorders, including fragile-X syndrome and autism (De Boulle et al. 1993; Bassell and Warren 2008; Hernandez et al. 2009). FMRP contains four highly conserved nucleic acid binding domains: three KH domains and an arginine-glycine rich domain (RGG). A number of studies determined that the RGG domain of FMRP specifically binds to G-rich mRNA sequences, some of which were shown to fold into unimolecular G4s (Brown et al. 2001; Darnell et al. 2001; Schaeffer et al. 2001; Phan et al. 2011). The G4s in these mRNAs were found to be polymorphic. Therefore, for structural studies, a conformationally uniform RNA G4 (*Sc1*) was selected in vitro against the full-length protein (Darnell et al. 2001; Ramos et al. 2003).

Solution and cocrystal structures of the RGG peptide complexed with

Sc1 (RGG–*Sc1*) revealed an RNA comprised of a three-tier G4, coaxial to an anti-parallel A-form duplex. The two are connected by a junctional mixed-sequence (i.e., not canonical, all-G) quartet (Fig. 3A; Phan et al. 2011; Vasilyev et al. 2015). The connectivity of the *Sc1* G4 was unprecedented and exhibits several noncanonical structural features (Fig. 3B). The two upper G-quartets (G11•G15•G20•G24 above G12•G16•G21•G25) assemble into a parallel G4 structure consisting of guanines in the *anti* conformation and backbones aligned in the same direction. Remarkably, the third G-quartet (G9•G6•G18•G26) is inverted relative to the top two, such that the strand polarity of the guanines of the top quartets is opposite that of the bottom (Fig. 3B). Three of the guanines of the third (inverted) G-quartet are not directly connected to the G-tracts, which differs from the classical G4 sequence motif (Patel et al. 2007; Agarwala et al. 2015; Malgowska et al. 2016; Lightfoot et al. 2019). These three guanines connect to the bottom G-tracts through mononucleotide loops, whereas G25 connects directly to G26. Furthermore, while the guanines adopt both C2′-*endo* and C3′-*endo* puckers, the majority are in the former (DNA-like) conformation.

The mixed-sequence quartet (U8•A17•U28•G29) that joins the G4 and duplex moieties of *Sc1* is held by a novel hydrogen-bond network (Fig. 3C). Two crystallographically well-ordered water molecules participate in inter-nucleobase hydrogen bonding. Other noncanonical tetrads have been observed in G4s, typically positioned at dimer interfaces or forming interactions with proteins (Cheong
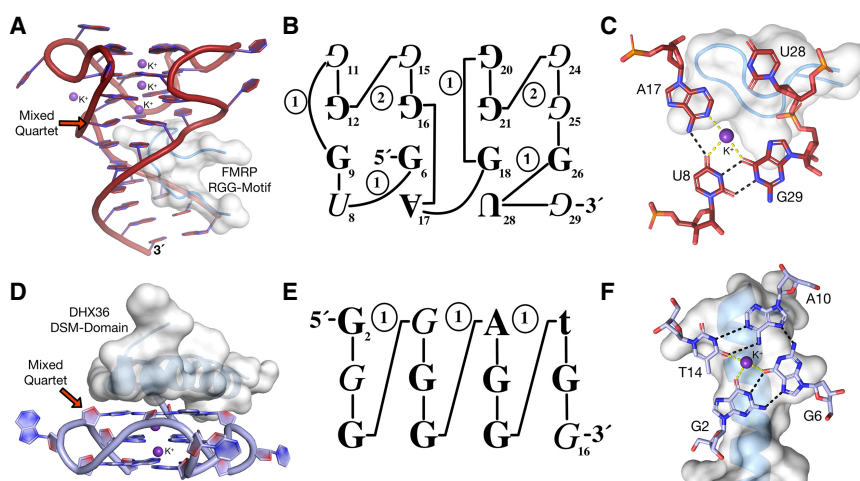

FIGURE 3. The Fragile-X peptide-aptamer RNA and DHX36-*c-Myc* complexes. (*A*) Cartoon representation of the three-dimensional structure of the RGG-*Sc1* complex (PDB:5DE5; Vasilyev et al. 2015). The location of the mixed-sequence quartet is indicated. The RGG peptide is highlighted by its translucent molecular surface. (*B*) Schematic of the *Sc1* quadruplex. (*C*) Hydrogen bond and coordination network of the mixed-sequence quartet of *Sc1*. The RGG peptide abuts this quartet. Red and purple spheres depict water molecules and K$^+$, respectively. (*D*) Cartoon representation of the *c-Myc* quadruplex bound to the amino-terminal α-helix (DSM) of DHX36 (PDB:5VHE; Chen et al. 2018). (*E*) Schematic of the *c-Myc* quadruplex bound to DHX36. (*F*) Detail of the mixed-sequence quartet.

and Moore 1992; Patel and Hosur 1999; Patel et al. 2000; Pan et al. 2003; Kimura et al. 2009; Chen et al. 2018; Andralojc et al. 2019). For instance, the solution structure of r(UGGUGGU) shows a tetramolecular G4 containing multiple U-tetrads. Of these, the 3′ terminal U-tetrad, contributes to the thermal stability of the G4, whereas in the DNA G4 of analogous sequence, the uridines lead to a 30°C lower T_m (Cheong and Moore 1992; Andralojc et al. 2019). The *Sc1* structure is stabilized by binding of the RGG peptide to a groove between its G4 and duplex moieties (Fig. 3A). In the peptide-free state, *Sc1* adopts multiple conformations, whereas binding of the RGG peptide selects for the G4 structure (Phan et al. 2011). In particular, the RGG peptide binding energy facilitates formation of the mixed, noncanonical tetrad.

The recent cocrystal structure of the DEAH helicase, DHX36, bound to the *c-Myc* DNA G4 is another example of protein binding-mediated stabilization of otherwise unstable mixed-sequence tetrads (Fig. 3D,E; Chen et al. 2018). DHX36 binding to the canonical *c-Myc* G4 (Fig. 2A) results in removal of the last guanine of the 3′-most G-stack from the quadruplex. The DNA rearranges such that two nucleotides previously in loops form part of the top tetrad of the three-tier quadruplex, whose composition is G•G•A•T (Fig. 3F). Differential scanning calorimetric analyses demonstrated that in the absence of the helicase, the G4 containing the G•G•A•T is much less stable than the canonical *c-Myc* G4 (Chen et al. 2018).

THE SPINACH FLUORESCENCE TURN-ON APTAMER

Fluorescence turn-on aptamers are RNAs selected in vitro to bind and activate otherwise weakly fluorescent small molecules, that is, conditional fluorophores (Bouhedda et al. 2017; Truong and Ferré-D'Amaré 2019). Their most common mechanism of action is to restrain the bound, photoexcited fluorophores in a planar conformation, suppressing nonradiative decay pathways (Trachman and Ferré-D'Amaré 2019). These RNA-fluorophore complexes have attracted great interest for their application as RNA analogs of fluorescent proteins and have been successfully used both, in vitro and in vivo for analysis of RNA folding and localization, as well as reporters for biosensors (Trachman and Ferré-D'Amaré 2019; Braselmann et al. 2020). Structures have been reported of multiple independently evolved fluorescence turn-on aptamers, which revealed diverse, and often unrelated three-dimensional architectures. Unexpectedly, many of these aptamers have fluorophore binding sites organized around G4 structural elements. It has been proposed that this may be a consequence of the mechanism of fluorescence activation used by these aptamers. The larger and flatter molecular surface of a G-quartet, compared to those of a base pair

or base triple, may facilitate restraining the photoexcited fluorophores (Warner et al. 2014).

Spinach is a fluorescence turn-on aptamer that induces by 1000-fold the fluorescence of 3,5-difluoro-4-hydroxy-benzylidene (DFHBI), a small-molecule analog of the intrinsic fluorophore of green fluorescent protein (Paige et al. 2011). Crystal structures of Spinach in complex with DFHBI revealed an RNA consisting of two A-form duplexes stacked coaxially on either side of a fluorophore binding site organized around a highly irregular, noncanonical quadruplex (Huang et al. 2014; Warner et al. 2014). Presence of a G4 in Spinach had not been predicted, as the RNA does not conform to the standard G4 sequence motif. In the Spinach-DFHBI complex structure, the fluorophore is held between a base triple and the top G-quartet (G26•G30•G65•G70) of the G4 motif (Fig. 4A). Subsequently, structures of the reselected variant aptamer *i*Spinach (Autour et al. 2016), which contains an identical G4 structure, in complex with two different fluorophores, have been reported (Fernandez-Millan et al. 2017; Jeng et al. 2021).

The RNA quadruplex motif of Spinach is among the most complex described to date. Two G-quartets stack onto a mixed-sequence quartet, (C28•U66•U73•G74), and the three tiers together coordinate two axial K⁺ ions (Fig. 4A,B). The connectivity of the G4 is a combination of parallel and antiparallel, and only four of the eight
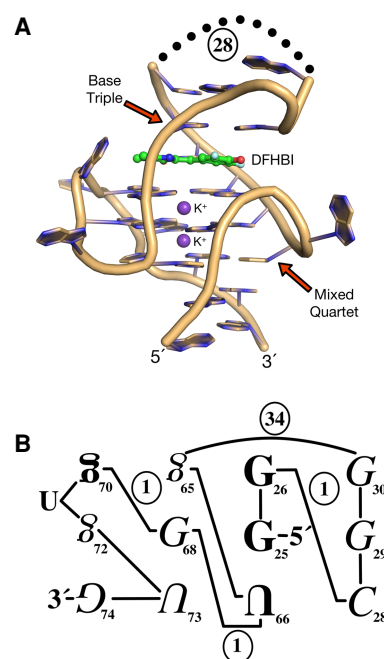

FIGURE 4. The Spinach fluorescence turn-on RNA aptamer. (*A*) Cartoon representation of the G4-containing core of Spinach bound to its cognate fluorophore DFHBI (PDB:4TS2; Warner et al. 2014). Locations of the mixed-sequence quartet and base triple are indicated. Dots indicate 28 nt omitted from the figure. (*B*) Schematic of the quadruplex element of Spinach.

guanines in the two G-quartets are sequential; the other four follow loops of various lengths. One of the loops (formally a diagonal loop) encompasses the base triple that stacks on the fluorophore as well as the entire A-form duplex that stacks coaxially on it. In the standard form of Spinach, it is 34 nt long. Before this long loop, the guanines are in the *anti* conformation and the connectivity is parallel. Following the loop, the connectivity is antiparallel, and the guanines deviate greatly from canonical RNA G4 stereochemistry. Three of these guanines adopt the *syn* conformation and only one is *anti*. Furthermore, all of these guanines adopt the 2′-*endo* ribose pucker. The three *syn* guanines are also in locally opposite polarity to the guanines 5′ of the 34 nt loop. The most notable difference of Spinach from canonical G4s is the drastic deviation from the classical sequence motif subsequent to the 34-nt loop. The guanines involved in the G-quartets are interrupted by nucleotides that form bulges or are incorporated into the mixed-sequence quartet. The mixed-sequence quartet is an integral part of the Spinach quadruplex, and beyond its composition, is rich in noncanonical features, including three nucleotides in inverted polarity and with 2′-*endo* puckers.

Analogous to the function of the mixed tetrad in the FMRP–*Sc1* complex (Phan et al. 2011; Vasilyev et al. 2015), the mixed tetrad in Spinach aids in the transition from the mixed G4 structure to an anti-parallel duplex. Unlike in the FMRP–*Sc1* complex, where the G4 stacks coaxially on a duplex only on one side, the Spinach G4 is in the middle of a continuous coaxial stack, being flanked by A-form duplexes on both sides. The elaborate, noncanonical connectivity of the Spinach G4 must arise, at least in part, as a solution to the problem of transitioning, without interrupting base stacking, from an antiparallel duplex, to a four-stranded G4, and back to an antiparallel duplex. Functionally, it has been proposed that the Spinach G4 is favorable for inducing fluorescence of its cognate fluorophore because G4s are planar and have a larger surface area than a base pair or base triple. In addition, it has been noted that electronic coupling between the fluorophore and the G4, as well as the coordinated axial cations, can modulate the photophysics of the turn-on aptamer-fluorophore complex (Warner et al. 2014; Trachman and Ferré-D'Amaré 2019).

THE CORN FLUORESCENCE TURN-ON APTAMER

To discover fluorescence turn-on aptamers with improved photophysical properties, several additional in vitro selection experiments have been reported (Shelke et al. 2018; Bouhedda et al. 2020; Braselmann et al. 2020). Selection against 3,5-difluoro-4-hydroxybenzyildene-imidazolinone-2-oxime (DFHO), a variant of DFHBI with additional conjugation, resulted in a small (36 nt), yellow fluorescent RNA, termed Corn (Song et al. 2017). The Corn-DFHO cocrystal structure (Warner et al. 2017) revealed a quasi-symmetric (see Jones and Ferré-D'Amaré 2015 for definition) homodimeric RNA, with one molecule of the fluorophore bound at the dimer interface (Fig. 5A). Biochemical studies of Corn demonstrated that the aptamer is a homodimer in solution (Warner et al. 2017). A crystal structure of the unliganded aptamer revealed a fully symmetric dimer with a collapsed fluorophore binding site. Structures of the Corn dimer in complex with the nonspecific fluorophores thioflavin T and thiazole orange also exhibited strict (crystallographic) symmetry, suggesting that local symmetry breaking at the interface is induced by DFHO binding (Sjekloća and Ferré-D'Amaré 2019).

Each of the Corn RNA chains forming the dimer (the two "protomers" in the sense of, e.g., Jones and Ferré-
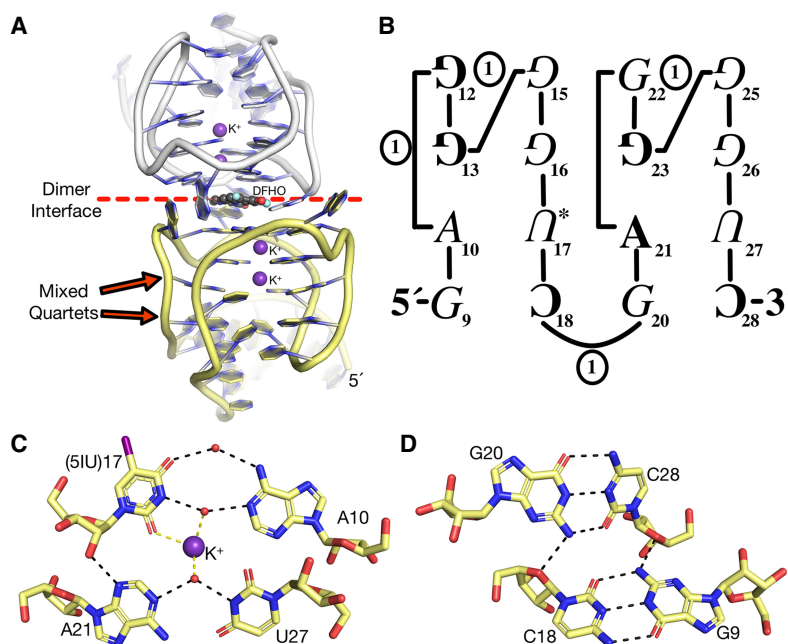

FIGURE 5. The Corn fluorescence turn-on aptamer. (*A*) Cartoon representation of the Corn homodimer with the fluorophore DFHO bound at the dimer interface (PDB:5BJO; Warner et al. 2017). Red arrows point to the two mixed-sequenced quartets. (*B*) Schematic of the quadruplex element of Corn. (*C*) Detail of the A•U•A•U quartet of Corn. Water molecules and K⁺ depicted as red and purple spheres, respectively. For structure determination U17 was replaced with 5-iodouracil (Warner et al. 2017). Subsequent Corn structures, containing RNA lacking halogens, showed the same conformation of the quartet (Sjekloća and Ferré-D'Amaré 2019). (*D*) Detail of the G•C•G•C quartet of Corn.

D'Amaré 2015) comprises a four-tiered mixed-sequence quadruplex, connected to an A-form duplex through an irregular junction. The Corn quadruplex contains two canonical G-quartets and two mixed-sequence quartets (Fig. 5B). The G-quartets of the quadruplex have parallel connectivity (except for one guanine that is locally inverted), all guanine bases are in the *anti* conformation, and most bases adopt a C3′-*endo* pucker. In contrast to the G-quartets, the mixed-sequence quartets have an anti-parallel connectivity, and unusual base-pairing interactions within each quartet (Fig. 5B). The mixed quartet adjacent to the G-quartets (A10•U17•A21•U27) exhibits noncanonical and water-mediated A•U pairing (Fig. 5C). These noncanonical interactions mediated by

FIGURE 6. The Mango family of fluorescence turn-on RNA aptamers. (*A*) Cartoon representation of the quadruplex portion of the cocrystal structure of Mango-I bound to the TO1-Biotin fluorophore (PDB: 5V3F; Trachman et al. 2017). (*B*) Schematic of Mango-I G4. (*C*) Schematic of the Mango-II G4 (PDB:6C63; Trachman et al. 2018). (*D*) Cartoon representation of the quadruplex portion of the cocrystal structure of Mango-III bound to TO1-biotin (PDB:6E8T; Trachman et al. 2019). (*E*) Schematic of the Mango-III G4. (*F*) Schematic of the Mango-IV G4 (PDB:6V9B; Trachman et al. 2020).

structured waters in mixed quartets have only been observed, to our knowledge, in the *Sc1* and Corn RNAs (Vasilyev et al. 2015; Warner et al. 2017). The next mixed-composition quartet (G9•C18•G20•C28) contains two G•C Watson–Crick base pairs, in which both base pairs display a pronounced negative roll (Fig. 5D; Olson et al. 2001) of ∼70°, causing the quartet to buckle. Previously, DNAs folding into a tetramolecular "fold-back" quadruplex have been observed adopting similar noncanonical G•C Watson–Crick base pairs that in turn pair into mixed quartets (Chu et al. 2018). Distal to this quartet, a junctional base triple joins the four-tier quadruplex of Corn with its A-form anti-parallel duplex.

THE MANGO FAMILY OF FLUORESCENCE TURN-ON APTAMERS

Selection experiments for small RNA turn-on aptamers that exhibit high affinity for the thiazole orange (TO) derivative, TO1-Biotin, yielded four aptamers, Mango-I, Mango-II, Mango-III, and Mango-IV (Dolgosheina et al. 2014; Autour et al. 2018; Trachman and Ferré-D'Amaré 2019). The sequences of these ∼30-nt long aptamers are related, and crystallographic structure determination (Trachman et al. 2017, 2018, 2019, 2020) confirmed that their ligand binding sites are all structured around G4s (Fig. 6). In terms of their overall structure, Mango-I and Mango-II (Fig. 6A–C) are most closely related. Both their structures are based on a three-tiered G4 that is linked to an A-form double helix by a GAAA-tetraloop-like junction. The sequence of Mango-IV is not highly divergent from those of Mango-I and Mango-II, but this RNA exists as a homodimer, and its G4 is stacked coaxially onto an A-form double helix
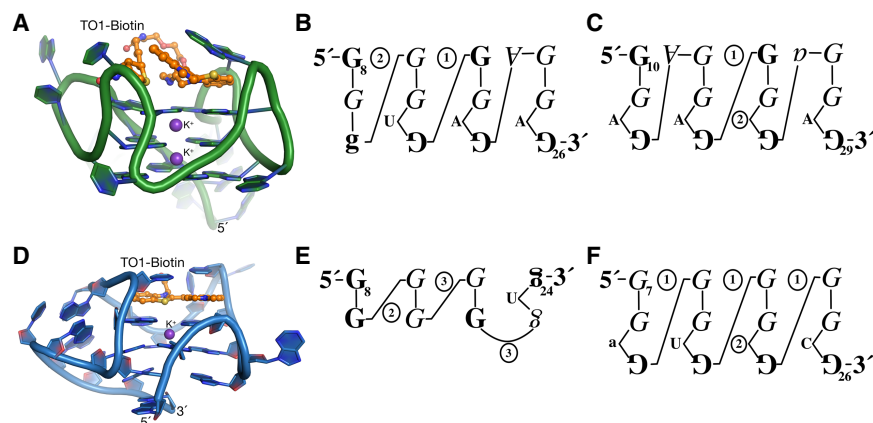
that links the two protomers through a domain-swapping interaction. Mango-III is the most divergent of the four aptamers, both in sequence and in structure. This monomeric aptamer consists of a two-tiered G4 stacking coaxially on a junctional base triple, which assists with the transition to a coaxially stacked A-form duplex. Unlike the other three Mango aptamers, the Mango-III structure incorporates a noncanonical helix formed between a propeller loop of the quadruplex and nucleotides 3′ of the G4, in a manner analogous to a pseudoknot (reminiscent in this regard to Mango-III, a solution NMR structure of a DNA G4 possessing a duplex hairpin located within a bulge was recently reported [Ngoc Nguyen et al. 2020]). Moreover, and unlike the TO1-Biotin binding sites of the other three Mango aptamers, which are open on the side opposite to the G4, the Mango-III fluorophore binding site is capped by a long-range, noncanonical Watson–Crick pair (Fig. 6A,D). The presence of tertiary interactions in Mango-III is consistent with it having, uniquely within aptamers of the family, multiphasic melting behavior.

The G4s of the Mango-I, Mango-II, and Mango-IV aptamer RNAs are all three-tiered, consisting of two parallel G-quartets stacked on top of a G-quartet with the opposite polarity (Fig. 6B,C,F). The G4 of Mango-III is only two-tiered, with its first two G-tracts being parallel, while the last one runs in the opposite direction (Fig. 6E). The guanines of the three-tiered G4s of Mango-I, Mango-II, and Mango-III are predominantly in the *anti* conformation, while their sugar puckers are both C2′-*endo* and C3′-*endo*. The Mango-I and Mango-II quadruplexes are most similar, in this regard, with the two outlying G-quartets having most riboses in the C2′-*endo* conformation and the middle G-quartet having all riboses in the

C3′-*endo* pucker. The G4 of Mango-IV differs in that its two parallel G-quartets are comprised of guanines with C3′-*endo* puckers exclusively. The puckers of the Mango-III G4 guanines is a mixture of C2′-*endo* and C3′-*endo*. For all four Mango aptamers, multiple bulges interrupt G-tracts forming the G-quartets, which are flanked by locally inverted guanosines except for Mango-III. Functionally, the loop and bulge nucleotides of the Mango aptamer G4s are very important, as they provide "flaps" that sequester their fluorophores against the flat surface of their quadruplexes (Fig. 6A). The bottom (locally inverted) G-quartet of Mango-I and Mango-II have expansions in the plane (i.e., into pentads and hexads) by adenosine nucleotides that originate in loops. Such expansions by loop nucleotides enlarge the surface area of G-quartets, and are commonly observed in dimeric G4 structures or interacting with small-molecules (Zhang et al. 2001; Liu et al. 2002; Mashima et al. 2009; Collie et al. 2011; Martadinata and Phan 2013).

OTHER RNA APTAMERS WITH COMPLEX G4 ELEMENTS

Spiegelmers (from the German word for mirror) are RNA aptamers made from the unnatural L-ribose nucleotides. The intrinsic resistance of spiegelmers to plasma exo- and endonucleases has made them of interest as novel therapeutics (Vater and Klussmann 2015). Complement is an essential component of the human innate immunity (Ricklin et al. 2010). Recognition of pathogens by complement causes the release of the anaphylatoxins, C3a and C5a, which function as signaling molecules (Zhou 2012). Elevated levels of C5a have been suggested to be involved in acute and chronic inflammation disorders. Substantial efforts have been made to develop novel inhibitors targeting complement and have been reviewed elsewhere (Ricklin et al. 2018). An in vitro selection experiment produced an L-DNA/L-RNA aptamer that binds C5a with high affinity, termed NOX-D20 (Hoehlig et al. 2013). The cocrystal structure of NOX-D20 in complex with C5a revealed a unique G4-containing nucleic acid (Fig. 7A; Yatime et al. 2015). Bound to C5a, NOX-D20 folds into a V-shaped structure comprised of a two-tier G4 and an anti-parallel duplex. Unusually, the cation coordinated by the NOX-D20 G-quartet is Ca^{2+}. Its two-tiered G4 structure has mixed connectivity and two bulges (Fig. 7B). Guanines in two of the G-stacks, which harbor bulges, have locally opposite polarity from the others. The guanines in this G4 are in both *anti* and *syn* conformations and adopt a mixture of C2′-*endo* and 3′-*endo* sugar puckers. The G4 structure contains two bulges, a 2 nt and a 4 nt, that interrupt two of the G-tracts. Interestingly, the tetranucleotide bulge forms a Watson–Crick and a noncanonical base pair with another loop region from the rest of the NOX-D20 structure.
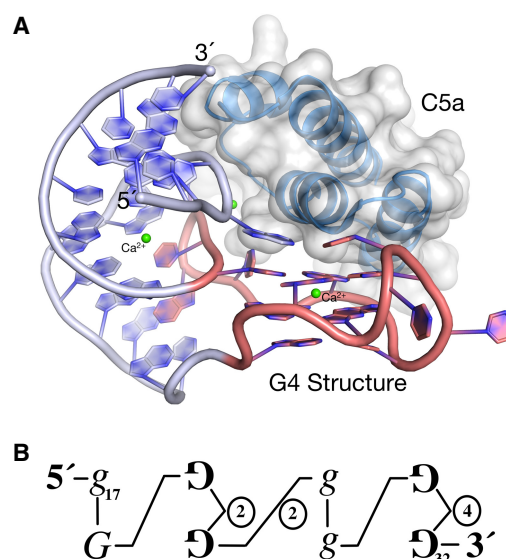

FIGURE 7. The NOX-D20 aptamer complex. (*A*) Cartoon representation of the structure of the NOX-D20 aptamer bound to C5a (PDB:4WB2; Yatime et al. 2015). The C5a four-helix bundle is highlighted by a translucent molecular surface. The antiparallel duplex and G4 structural elements of NOX-D20 are represented in white and salmon, respectively. Green spheres denote Ca^{2+}. (*B*) Schematic of the NOX-D20 quadruplex.

These are important in maintaining the overall fold of the nucleic acid (Yatime et al. 2015).

Prion proteins (PrP) are proteinaceous infectious particles, and the etiological agent for multiple neurodegenerative diseases that include bovine spongiform encephalopathy, scrapie of sheep, and Creutzfeldt-Jakob disease (Prusiner 1998). Several aptamers have been produced by in vitro selection to preferentially bind with high affinity to isoforms of PrP (Weiss et al. 1997; Proske et al. 2002; Rhie et al. 2003; Sekiya et al. 2005; Nishikawa et al. 2007; Murakami et al. 2008), some of which are potential therapeutic agents (Proske et al. 2002; Rhie et al. 2003; Mashima et al. 2013;). Of these in vitro selected aptamers, the NMR solution structure of the sequence r(GGAGGAGGAGGA) (R12) was shown to fold into a two-tiered G4 that forms a "tail-to-tail" homodimer in solution (Fig. 8A; Mashima et al. 2009, 2013). Recently, another NMR solution structure was published of an RNA aptamer that linked two R12 sequences together in tandem, which resembled the R12 structure (Mashima et al. 2020). In the peptide-bound R12 complex structure, the peptide directly interacts with the solvent exposed face of the G-quartet and loop nucleotides. In comparison to the other G4 RNA aptamers discussed above, the R12 aptamer has the simplest structure and mostly conforms to canonical RNA G4 stereochemistry. R12 nucleotides adopt the *anti* conformation and a parallel connectivity but have predominantly C2′-*endo* puckers (Fig. 8B). The G-quartet positioned at the dimer interface forms a hexad
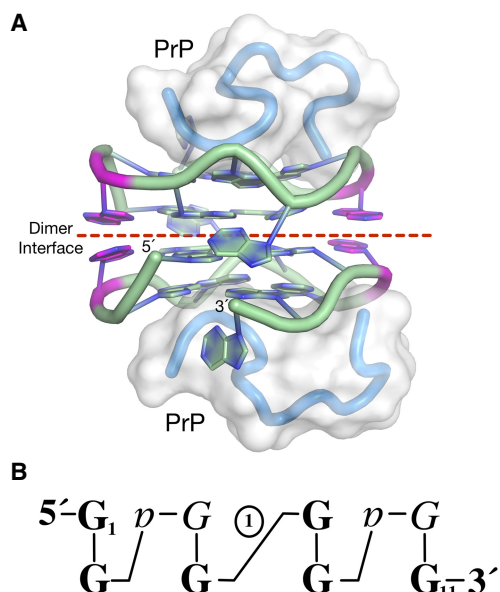
FIGURE 8. PrP Aptamer. (*A*) Cartoon representation of the three-dimensional structure of an anti-PrP aptamer bound to a small peptide from the soluble isoform of the protein (PDB:2RSK; Mashima et al. 2013). The PrP-derived peptide is highlighted by its molecular surface. Dashed red line denotes the dimer interface of the aptamer. Adenosines that extend the interfacial G-quartet into a hexad are colored magenta. (*B*) Schematic of the PrP aptamer illustrating the canonical RNA G4 stereochemistry.

by addition of two loop adenosines, which extend the dimer interface. Its NMR solution structure shows that the DNA counterpart of R12 forms a heptad at the dimer interface, which was proposed to result from differences in sugar pucker (Matsugami et al. 2001; Mashima et al. 2009).

THE STRUCTURAL SYNTAX OF RNA G-QUADRUPLEXES

The recent structure determinations of elaborate RNA G4s show that, once emancipated from the requirement that G-stacks be assembled from consecutive guanines, the structural possibilities of RNA G4s are unshackled. (A summary of structural parameters observed in the RNA aptamers discussed can be found in Table 2.) The structures described above highlight that by expanding their sequence complexity, RNA G-quartets can overcome the stereochemical constraints imposed by the ribose 2′-OH and have additional conformational freedom. For instance, TERRA possesses low sequence complexity (Fig. 2D) resulting in a G4 structure that conforms to the canonical RNA G4 stereochemistry, whereas the Spinach turn-on aptamer (Fig. 4B) is comprised of multiple bulges and a 34-nt. diagonal loop, which support the intricate architecture of its quadruplex.

The results of genome-wide sequencing suggest that bulges are the most abundant noncanonical structural feature in putative G4 sequences (Chambers et al. 2015). In most of the elaborate RNA G4 structures discussed above, bulges of varying lengths interrupt their G-tracts. Comparison of these bulged nucleotides suggests that they prefer to adopt the *anti* conformation and the C2′-*endo* sugar pucker. Thus far, only the Mango-IV (Fig. 6F) and NOX-D20 (Fig. 7B) structures have either *syn* or C3′-*endo* bulged nucleotides. Bulges present in DNA and RNA G4 structures have expanded ranges in backbone torsion angles (Meier et al. 2020), which would contribute to overcoming the steric barriers imposed by the ribose 2′-OH. The recent RNA G4 structures suggest that inversion of the strand polarity of adjacent guanines (within a G-stack) requires the presence of intervening bulges or loops. This is apparent in the inverted G-quartets of, Mango-I, Mango-II, and Mango-IV, where bulges precede the local inversion of strand polarity (Fig. 6B,C,F). Spinach (Fig. 4B) also has a bulge and a long loop prior to a guanine with inverted polarity. Interestingly, the RGG-motif–Sc1 complex (Fig. 3B), Corn turn-on aptamer (Fig. 5B), and NOX-D20–C5a complex structures (Fig. 7B) have guanines with inverted polarity immediately before a loop or bulge. In these structures, some of the inverted guanines are directly connected. In all cases, inversion in strand polarity is accompanied by a change in sugar pucker. Nucleotides in the *syn* conformation also appear to exist preferentially after a bulge or loop.

Another noncanonical structural feature exhibited by the some of the RNA G4s are inclusions (pentads, hexads, etc.) of nucleotides in the plane of quartets. Quartets with inclusions are observed forming at dimer interfaces of two stacked protomers (Kolesnikova and Curtis 2019; Lightfoot et al. 2019), which is also demonstrated by the RNA G4s. Inclusions were shown to be indispensable for the formation of the homodimer of d(GGA)₄ (Matsugami et al. 2001). Comparison of the RNA G4s with inclusions shows that the incorporated nucleotide in the quartet plane prefers the 2′-*endo* sugar pucker but is agnostic to adopting the *anti* or *syn* conformation. Additionally, inclusion of a nucleotide also appears to affect the sugar pucker of the subsequent guanine in the G4 structure. As shown by Mango-I (Fig. 6B), Mango-II (Fig. 6C) and the anti-PrP aptamer (Fig. 8B), the guanine following the inclusion prefers to be *anti* and 2′-*endo*.

Noncanonical quartets adjacent to G-quartets have now been observed in several structures (e.g., RGG–Sc1, Fig. 3; Spinach, Fig. 4; Corn, Fig. 5). These appear to be integral to the structures of these aptamers, serving to connect the G-quadruplexes to other structural elements of the RNAs, as do base triples (e.g., in the structures of Spinach and Mango-III). Although much of the structural work on G4s up to the present has concentrated on isolated quadruplexes, it is likely that many

TABLE 2. Summary of structural features of representative simple and elaborate G-quadruplexes

Name (PDB ID)	Pentose	Connectivity	anti/syn[a]	C2'-endo/C3'-endo[a]	Up/Down nucleotides[b]	Non-G quartets	Bulges/inclusions	Total no. of nucleotides[c]
c-Myc (1XAV)	DNA	Parallel	12/0	10/2	12/0	-	-	16
TERRA (3IBK)	RNA	Parallel	12/0	5/7	12/0	-	-	18 (9 × 2)
RGG-*Sc1* (5DE5)	RNA	Mixed	16/0	10/6	5/11	1	2/0	22
c-Myc/DHX36 (5VHE)	DNA	Parallel	11/1	9/3	12/0	1	-	15
Spinach (4TS2)	RNA	Mixed	9/3	4/8	6/6	1	3/0	49
Corn (5BJO)	RNA	Mixed	16/0	6/10	5/11	2	-	36
Mango-I (5V3F)	RNA	Parallel	11/1	6/6	9/3	-	3/1	31
Mango-II (6C63)	RNA	Parallel	12/0	5/7	8/4	-	4/2	36
Mango-III (6E8T)	RNA	Mixed	6/2	4/4	6/2	-	1/0	17
Mango-IV (6V9B)	RNA	Parallel	12/0	4/8	8/4	-	4/0	28
NOX-D20/C5a (4WB2)	Mixed	Parallel	5/3	4/4	4/4	-	2/0	40
Anti-PrP (2RSK)	RNA	Parallel	8/0	6/2	8/0	-	0/2	12

[a]Number of nucleotides that are part of quartets, with indicated stereochemistry. Nucleotides in loops or bulges are excluded.
[b]Number of nucleotides with the same or inverted polarity (Up and Down, respectively) as the 5' G-quartet nucleotide. Nucleotides in loops or bulges are excluded.
[c]Total number of nucleotides between the 5' and 3' nucleotides that are part of quartets (G- quartet or noncanonical), including loop and bulge nucleotides.

quadruplexes will be part of more elaborate structures, and as such, the structural solutions to transitioning from a four-stranded structure to other nucleic acids structures (and in particular the most common structure, the A-form duplex) will reveal intricate and idiosyncratic arrangements of nucleotides, not necessarily limited to guanines. Based on recent structures, we suggest that the apparent stereochemical simplicity of G4s, and specifically of RNA G4s, is probably a misrepresentation arising from the difficulty in preparing biochemically tractable samples. This may also resolve the apparent paradox that G4s appear to be more common among in vitro selected than in natural RNAs. That is, cellular G4s may have evolved to be energetically metastable, and therefore challenging to characterize structurally. Consistent with this is the existence of many helicase proteins (Mendoza et al. 2016) capable of recognizing and destabilizing G4s. The future study of functional nucleic acids, either natural or artificial, that incorporate quadruplexes is likely to increase our understanding of the complex structural syntax of RNA G-quadruplexes, and their interplay with other RNA structural elements.

REFERENCES

Agarwal T, Jayaraj G, Prakash Pandey S, Agarwala P, Maiti S. 2012. RNA G-quadruplexes: G-quadruplexes with "U" turns. *Curr Pharm Design* **18:** 2102–2111. doi:10.2174/138161212799 958468

Agarwala P, Pandey S, Maiti S. 2015. The tale of RNA G-quadruplex. *Org Biomol Chem* **13:** 5570–5585. doi:10.1039/C4OB02681K

Ambrus A, Chen D, Dai J, Jones RA, Yang D. 2005. Solution structure of the biologically relevant G-quadruplex element in the human c-Myc promoter. Implications for G-quadruplex stabilization. *Biochemistry* **44:** 2048–2058. doi:10.1021/bi048242p

Andralojc W, Malgowska M, Sarzynska J, Pasternak K, Szpotkowksi K, Kierzek R, Gdaniec Z. 2019. Unraveling the structural basis for the exceptional stability of RNA G-quadruplexes capped by a uridine tetrad at the 3' terminus. *RNA* **25:** 121–134. doi:10.1261/rna.068163.118

Arora A, Maiti S. 2009. Differential biophysical behavior of human telomeric RNA and DNA quadruplex. *J Phys Chem B* **113:** 10515–10520. doi:10.1021/jp810638n

Arora A, Suess B. 2011. An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation. *RNA Biol* **8:** 802–805. doi:10.4161/rna.8.5.16038

Autour A, Westhof E, Ryckelynck M. 2016. iSpinach: a fluorogenic RNA aptamer optimized for in vitro applications. *Nucleic Acids Res* **44:** 2491–2500. doi:10.1093/nar/gkw083

Autour A, Jeng SC, Cawte AD, Abdolahzadeh A, Galli A, Panchapakesan SS, Rueda D, Ryckelynck M, Unrau PJ. 2018. Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat Commun* **9:** 656. doi:10.1038/s41467-018-02993-8

Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J. 2007. Telomeric repeat-containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318:** 798–801. doi:10.1126/science.1147182

Bang I. 1910. Untersuchungen über die Guanylsäure. *Biochemische Zeitschr* **26:** 293–311.

Bassell GJ, Warren ST. 2008. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* **60:** 201–214. doi:10.1016/j.neuron.2008.10.004

Bhattcharyya D, Arachchilage GM, Basu S. 2016. Metal cations in G-quadruplex folding and stability. *Front Chem* **4:** 38. doi:10.3389/fchem.2016.00038

Bouhedda F, Autour A, Ryckelynck M. 2017. Light-up RNA aptamers and their cognate fluorogens: from their development to their applications. *Int J Mol Sci* **19:** 44. doi:10.3390/ijms19010044

Bouhedda F, Fam KT, Collot M, Autour A, Marzi S, Klymchenko A, Ryckelynck M. 2020. A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nat Chem Biol* **16:** 69–76. doi:10.1038/s41589-019-0381-8

Braselmann E, Rathbun C, Richards EM, Palmer AE. 2020. Illuminating RNA biology: tools for imaging RNA in live mammalian cells. *Cell Chem Biol* **27:** 891–903. doi:10.1016/j.chembiol.2020.06.010

Brown V, Jin P, Ceman S, Darnell JC, O'Donell WT, Tenenbaum SA, Jin X, Feng Y, Wilkinson KD, Keene JD. 2001. Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107:** 477–487. doi:10.1016/S0092-8674(01)00568-2

Calabrese DR, Chen X, Leon EC, Gaikwad SM, Phyo Z, Hewitt WM, Alden S, Hilimire TA, He F, Michalowski AM. 2018. Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-quadruplex. *Nat Commun* **9:** 4229. doi:10.1038/s41467-018-06315-w

Carvalho J, Mergny J-L, Salgado GF, Queiroz JA, Cruz C. 2020. G-quadruplex, friend or foe: the role of the G-quartet in anticancer strategies. *Trends Mol Med* **26:** 848–861. doi:10.1016/j.molmed.2020.05.002

Chambers VS, Marisco G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. 2015. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol* **33:** 877–881. doi:10.1038/nbt.3295

Chen MC, Tippana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S, Ferré-D'Amaré AR. 2018. Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **558:** 465–469. doi:10.1038/s41586-018-0209-9

Cheong C, Moore PB. 1992. Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures. *Biochemistry* **31:** 8406–8414. doi:10.1021/bi00151a003

Chu B, Zhang D, Hwang W, Paukstelis PJ. 2018. Crystal structure of a tetrameric DNA fold-back quadruplex. *J Am Chem Soc* **140:** 16291–16298. doi:10.1021/jacs.8b10153

Chung WJ, Heddi B, Schmitt E, Lim KW, Mechulam Y, Phan AT. 2015. Structure of a left-handed DNA G-quadruplex. *Proc Natl Acad Sci* **112:** 2729–2733. doi:10.1073/pnas.1418718112

Collie GW, Haider SM, Neidle S, Parkinson GN. 2010. A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res* **38:** 5569–5580. doi:10.1093/nar/gkq259

Collie GW, Sparapani S, Parkinson GN, Neidle S. 2011. Structural basis of telomeric RNA quadruplex-acridine ligand recognition. *J Am Chem Soc* **133:** 2721–2728. doi:10.1021/ja109767y

Dang CV. 2012. MYC on the path to cancer. *Cell* **149:** 22–35. doi:10.1016/j.cell.2012.03.003

Darnell JC, Jensen KB, Jin P, Brown V, Warren ST, Darnell RB. 2001. Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* **107:** 489–499. doi:10.1016/S0092-8674(01)00566-9

De Boulle K, Verkerk AJ, Reyniers E, Vits L, Hendrickx J, Van Roy B, Van den Bos F, de Graaff E, Oostra BA, Willems PJ. 1993. A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat Genet* **3:** 31–35. doi:10.1038/ng0193-31

Dolgosheina EV, Jeng SC, Panchapakesan SS, Cojocaru R, Chen PS, Wilson PD, Hawkins N, Wiggins PA, Unrau PJ. 2014. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem Biol* **9:** 2412–2420. doi:10.1021/cb500499x

Fay MM, Lyons SM, Ivanov P. 2017. RNA G-quadruplexes in biology: principles and molecular mechanisms. *J Mol Biol* **429:** 2127–2147. doi:10.1016/j.jmb.2017.05.017

Fernandez-Millan P, Autour A, Ennifar E, Westhof E, Ryckelynck M. 2017. Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI. *RNA* **23:** 1788–1795. doi:10.1261/rna.063008.117

Gellert M, Lipsett MN, Davies DR. 1962. Helix formation by guanylic acid. *Proc Natl Acad Sci* **48:** 2013–2018. doi:10.1073/pnas.48.12.2013

Hänsel-Hertsch R, Di Antonio M, Balasubramanian S. 2017. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* **18:** 279–284.

Hernandez RN, Feinberg RL, Vaurio R, Passanante NM, Thompson RE, Kaufmann WE. 2009. Autism spectrum disorder in fragile X syndrome: a longitudinal evaluation. *Am J Med Genet A* **149:** 1125–1137. doi:10.1002/ajmg.a.32848

Hoehlig K, Maasch C, Shushakova N, Buchner K, Huber-Lang M, Purschke WG, Vater A, Klussmann S. 2013. A novel C5a-neutralizing mirror-image (l-)aptamer prevents organ failure and improves survival in experimental sepsis. *Mol Ther* **21:** 2236–2246. doi:10.1038/mt.2013.178

Huang H, Suslov NB, Li N, Shelke SA, Evans ME, Koldobskaya Y, Rice PA, Piccirilli JA. 2014. A G-quadruplex–containing RNA activates fluorescence in a GFP-like fluorophore. *Nat Chem Biol* **10:** 686–691. doi:10.1038/nchembio.1561

Hurley LH, Von Hoff DD, Siddiqui-Jain A, Yang D. 2006. Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Semin Oncol* **33:** 498–512. doi:10.1053/j.seminoncol.2006.04.012

Jeng SCY, Trachman RJ III, Weissenboeck F, Truong L, Link KA, Jepsen MDE, Knutson JR, Andersen ES, Ferré-D'Amaré AR, Unrau PJ. 2021. Fluorogenic aptamers resolve the flexibility of RNA junctions using orientation-dependent FRET. *RNA* **27:** 433–444 (this issue). doi:10.1261/rna.078220.120

Jhunjhunwala A, Ali Z, Bhattacharya S, Halder A, Mitra A, Sharma P. 2020. A comprehensive survey on the nature of ring:ring nucleobase stacking interactions in RNA: occurrence, structural variability and classification of the associated contacts. bioRxiv doi:10.1101/2020.02.28.970798

Joachimi A, Benz A, Hartig JS. 2009. A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg Med Chem* **17:** 6811–6815. doi:10.1016/j.bmc.2009.08.043

Jones CP, Ferré-D'Amaré AR. 2015. RNA quaternary structure and global symmetry. *Trends Biochem Sci* **40:** 211–220. doi:10.1016/j.tibs.2015.02.004

Kimura T, Xu Y, Komiyama M. 2009. Human telomeric RNA r(UAGGGU) sequence forms parallel tetraplex structure with U-quartet. *Nucleic Acids Symp Ser* **53:** 239–240. doi:10.1093/nass/nrp120

Kolesnikova S, Curtis EA. 2019. Structure and function of multimeric G-quadruplexes. *Molecules* **24:** 3074. doi:10.3390/molecules24173074

Kumari S, Bugaut A, Huppert JL, Balasubramanian S. 2007. An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nat Chem Biol* **3:** 218–221. doi:10.1038/nchembio864

Kwok CK, Marisco G, Sahakyan AB, Chambers VS, Balasubramanian S. 2016. rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat Methods* **13:** 841–844.

Lee DSM, Ghanem LR, Barash Y. 2020. Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. *Nat Commun* **11:** 527. doi:10.1038/s41467-020-14404-y

Leontis NB, Westhof E. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA* **7:** 499–512. doi:10.1017/S1355838201002515

Lightfoot HL, Hagen T, Tatum NJ, Hall J. 2019. The diverse structural landscape of quadruplexes. *FEBS Lett* **593:** 2083–2102. doi:10.1002/1873-3468.13547

Liu H, Matsugami A, Katahira M, Uesugi S. 2002. A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops. *J Mol Biol* **322:** 955–970. doi:10.1016/S0022-2836(02)00876-8

Malgowska M, Czajczynska K, Gudanis D, T'worak A, Gdaniec Z. 2016. Overview of the RNA G-quadruplex structures. *Acta Biochim Pol* **63:** 609–621. doi:10.18388/abp.2016_1335

Martadinata H, Phan AT. 2009. Structure of propeller-type parallel-standed RNA G-quadruplexes, formed by human telomeric RNA sequences in K+ solution. *J Am Chem Soc* **131:** 2570–2578. doi:10.1021/ja806592z

Martadinata H, Phan AT. 2013. Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K+ solution. *Biochemistry* **52:** 2176–2183. doi:10.1021/bi301606u

Mashima T, Matsugami A, Nishikawa F, Nishikawa S, Katahira M. 2009. Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein. *Nucleic Acids Res* **37:** 6249–6258. doi:10.1093/nar/gkp647

Mashima T, Nishikawa F, Kamatari YO, Fujiwara H, Saimura M, Nagata T, Kodoki T, Nishikawa S, Kuwata K, Katahira M. 2013. Anti-prion activity of an RNA aptamer and its structural basis. *Nucleic Acids Res* **41:** 1355–1362. doi:10.1093/nar/gks1132

Mashima T, Lee JH, Kamatari YO, Hayashi T, Nagata T, Nishikawa F, Nishikawa S, Kinoshita M, Kuwata K, Katahira M. 2020. Development and structural determination of an anti-PrP^C aptamer that blocks pathological conformational conversion of prion protein. *Sci Rep* **10:** 4934. doi:10.1038/s41598-020-61966-4

Matsugami A, Ouhashi K, Kanagawa M, Liu H, Kanagawa S, Uesugi S, Katahira M. 2001. An intramolecular quadruplex of (GGA) (4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. *J Mol Biol* **313:** 255–269. doi:10.1006/jmbi.2001.5047

McRae EKS, Booy EP, Padilla-Meier GP, McKenna SA. 2017. On characterizing the interactions between proteins and guanine quadruplex structures of nucleic acids. *J Nucleic Acids* **2017:** 9675348. doi:10.1155/2017/9675348

Meier M, Moya-Torres A, Krahn NJ, McDougall M, Orris GL, McRae EKS, Booy EP, McEleney K, Patel TR, McKenna SA, et al. 2020. Structure and hydrodynamics of a DNA G-quadruplex with a cytosine bulge. *Nucleic Acids Res* **46:** 5319–5331. doi:10.1093/nar/gky307

Mendoza O, Bourdoncle A, Boulé JB, Brosh RMJ, Mergny JL. 2016. G-quadruplexes and helicases. *Nuclec Acids Res* **44:** 1989–2006. doi:10.1093/nar/gkw079

Murakami K, Nishikawa F, Noda K, Yokoyama T, Nishikawa S. 2008. Anti-bovine prion protein RNA aptamer containing tandem GGA repeat interacts both with recombinant bovine prion protein and its β isoform with high affinity. *Prion* **2:** 73–80. doi:10.4161/pri.2.2.7024

Murat P, Marisco G, Herdy B, Ghanbarian AT, Portella G, Balasubramanian S. 2018. RNA G-quadruplexes at upstream open reading frames cause DHX36- and DHX9-dependent translation of human mRNAs. *Genome Biol* **19:** 229. doi:10.1186/s13059-018-1602-2

Murray LJW, Arendall WB, Richardson DC, Richardson JS. 2003. RNA backbone is rotameric. *Proc Natl Acad Sci* **100:** 13904–13909. doi:10.1073/pnas.1835769100

Neidle S. 2017. Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat Rev Chem* **1:** 41. doi:10.1038/s41570-017-0041

Ngoc Nguyen TQ, Lim KW, Phan AT. 2020. Duplex formation in a G-quadruplex bulge. *Nucleic Acids Res* **48:** 10567–10575. doi:10.1093/nar/gkaa738

Nishikawa F, Murakami K, Noda K, Yokoyama T, Nishikawa S. 2007. Detection of structural changes of RNA aptamer containing GGA repeats under the ionic condition using the microchip electrophoresis. *Nucleic Acids Symp Ser* **51:** 397–398. doi:10.1093/nass/nrm199

Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, et al. 2001. A standard reference frame for the description of nucleic acid base-pair geometry. *J Mol Biol* **313:** 229–237. doi:10.1006/jmbi.2001.4987

Paige JS, Wu KY, Jaffrey SR. 2011. RNA mimics of green fluorescent protein. *Science* **333:** 642–646. doi:10.1126/science.1207339

Pan B, Xiong Y, Shi K, Deng J, Sundralingam M. 2003. Crystal structure of an RNA purine-rich tetraplex containing adenine tetrads: implications for specific binding in RNA tetraplexes. *Structure* **11:** 815–823. doi:10.1016/S0969-2126(03)00107-2

Patel PK, Hosur RV. 1999. NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by *Saccharomyces cerevisiae* telomere repeats. *Nucleic Acids Res* **27:** 2457–2464. doi:10.1093/nar/27.12.2457

Patel PK, Bhavesh NS, Hosur RV. 2000. NMR observation of a novel C-tetrad in the structure of the SV40 repeat sequence GGGCGG. *Biochem Biophys Res Commun* **270:** 967–971. doi:10.1006/bbrc.2000.2479

Patel DJ, Phan AT, Kuryavyi V. 2007. Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher-order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res* **35:** 7429–7455. doi:10.1093/nar/gkm711

Phan AT, Kuryavyi V, Darnell JC, Serganov A, Majumdar A, Ilin S, Raslin T, Polonskaia A, Chen C, Clain D, et al. 2011. Structurure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat Struct Mol Biol* **18:** 796–804. doi:10.1038/nsmb.2064

Proske D, Gilch S, Wopfner F, Schätzl HM, Winnacker EL, Famulok M. 2002. Prion-protein-specific aptamer reduces PrPSc formation. *ChemBioChem* **3:** 717–725. doi:10.1002/1439-7633(20020802)3:8<717::AID-CBIC717>3.0.CO;2-C

Prusiner SB. 1998. Prions. *Proc Natl Acad Sci* **95:** 13363–13383. doi:10.1073/pnas.95.23.13363

Ramos A, Hollingworth D, Pastore A. 2003. G-quartet-dependent recognition between the FMRP RGG box and RNA. *RNA* **9:** 1198–1207. doi:10.1261/rna.5960503

Rhie A, Kirby L, Sayer N, Wellesley R, Disterer P, Sylvester I, Gill A, Hope J, James W, T'ahiri-Alaoui A. 2003. Characterization of 2′-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. *J Biol Chem* **278:** 39697–39705. doi:10.1074/jbc.M305297200

Ricklin D, Hajishengallis G, Yang K, Lambris JD. 2010. Complement: a key system for immune surveillance and homeostasis. *Nat Immunol* **11:** 785–797. doi:10.1038/ni.1923

Ricklin D, Mastellos DC, Reis ES, Lambris JD. 2018. The renaissance of complement therapeutics. *Nat Rev Nephrol* **14:** 26–47. doi:10.1038/nrneph.2017.156

Saenger W. 1984. *Principles of nucleic acid structure*. Springer-Verlag, New York.

Saranathan N, Vivekanandan P. 2019. G-quadruplexes: more than just a kink in microbial genomes. *Trends Microbiol* **27:** 148–163. doi:10.1016/j.tim.2018.08.011

Sauer M, Paeschke K. 2017. G-quadruplex unwinding helicases and their function in vivo. *Biochem Soc Trans* **45:** 1173–1182. doi:10.1042/BST20170097

Sauer M, Juranek SA, Marks J, De Magis A, Kazemier HG, Hilbig D, Benhalevy D, Wang X, Hafner M, Paeschke K. 2019. DHX36 prevents the accumulation of translationally inactive mRNAs with G4-structures in untranslated regions. *Nat Commun* **10:** 2421. doi:10.1038/s41467-019-10432-5

Schaeffer C, Bardoni B, Mandel JL, Ehresmann B, Ehresmann C, Moine H. 2001. The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J* **20:** 4803–4813. doi:10.1093/emboj/20.17.4803

Sekiya S, Nishikawa F, Noda K, Kumar P, Yokoyama T, Nishikawa S. 2005. In vitro selection of RNA aptamers against cellular and abnormal isoform of mouse prion protein. *Nucleic Acids Symp Ser* **49:** 361–362.

Shelke SA, Shao Y, Laski A, Koirala D, Weissman BP, Fuller JR, Tan X, Constantin TP, Waggoner AS, Bruchez MP, et al. 2018. Structural basis for activation of fluorogenic dyes by an RNA aptamer lacking a G-quadruplex motif. *Nat Commun* **9:** 4542. doi:10.1038/s41467-018-06942-3

Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. 2002. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci* **99:** 11593–11598. doi:10.1073/pnas.182256799

Sissi C, Gatto B, Palumbo M. 2011. The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective. *Biochimie* **93:** 1219–1230. doi:10.1016/j.biochi.2011.04.018

Sjekloća L, Ferré-D'Amaré AR. 2019. Binding between G-quadruplexes at the homodimer interface of the Corn RNA aptamer strongly activates thioflavin T fluorescence. *Cell Chem Biol* **26:** 1159–1168. doi:10.1016/j.chembiol.2019.04.012

Song W, Filonov GS, Kim H, Hirsch M, Li X, Moon JD, Jaffrey SR. 2017. Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat Chem Biol* **13:** 1187–1194. doi:10.1038/nchembio.2477

Trachman RJ III, Ferré-D'Amaré AR. 2019. Tracking RNA with light: selection, structure and design of fluorescence turn-on RNA aptamers. *Q Rev Biophys* **52:** e8. doi:10.1017/S0033583519000064

Trachman RJ III, Demeshkina NA, Lau MWL, Pahchapakesan SSS, Jeng SCY, Unrau PJ, Ferré-D'Amaré AR. 2017. Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat Chem Biol* **13:** 807–813. doi:10.1038/nchembio.2392

Trachman RJ III, Abdolahzadeh A, Andreoni A, Cojocaru R, Knutson JR, Ryckelynck M, Unrau PJ, Ferré-D'Amaré AR. 2018. Crystal structures of the Mango-II RNA aptamer reveal heterogeneous fluorophore binding and guide engineering of variants with improved selectivity and brightness. *Biochemistry* **57:** 3544–3548. doi:10.1021/acs.biochem.8b00399

Trachman RJ III, Autour A, Jeng SCY, Abdolahzadeh A, Andreoni A, Cojocaru R, Garipov R, Dolgosheina EV, Knutson JR,

Ryckelynck M, et al. 2019. Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nat Chem Biol* **15:** 472–479. doi:10.1038/s41589-019-0267-9

Trachman RJ III, Cojocaru R, Wu D, Piszczek G, Ryckelynck M, Unrau PJ, Ferré-D'Amaré AR. 2020. Structure-guided engineering of the homodimeric Mango-IV fluorescence turn-on aptamer yields an RNA FRET pair. *Structure* **28:** 776–785. doi:10.1016/j.str.2020.04.007

Truong L, Ferré-D'Amaré AR. 2019. From fluorescent proteins to fluorogenic RNAs: tools for imaging cellular macromolecules. *Protein Sci* **28:** 1374–1386. doi:10.1002/pro.3632

Varshney D, Spiegel J, Zyner K, Tannahill D, Balasubrmanian S. 2020. The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **21:** 459–474. doi:10.1038/s41580-020-0236-x

Vasilyev N, Polonskaia A, Darnell JC, Darnell RB, Patel DJ, Serganov A. 2015. Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proc Natl Acad Sci* **112:** E5391–E5400. doi:10.1073/pnas.1515737112

Vater A, Klussmann S. 2015. Turning mirror-image oligonucleotides into drugs: the evolution of Spiegelmer® therapeutics. *Drug Discov Today* **20:** 147–155. doi:10.1016/j.drudis.2014.09.004

Warner KD, Chen MC, Song W, Strack RL, Thorn A, Jaffrey SR, Ferré-D'Amaré AR. 2014. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat Struct Mol Biol* **21:** 658–663. doi:10.1038/nsmb.2865

Warner KD, Sjekloća L, Song W, Filonov GS, Jaffrey SR, Ferré-D'Amaré AR. 2017. A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat Chem Biol* **13:** 1195–1201. doi:10.1038/nchembio.2475

Weiss S, Proske D, Neumann M, Groschup MH, Kretzschmar HA, Famulok M, Winnacker EL. 1997. RNA aptamers specifically interact with the prion protein PrP. *J Virol* **71:** 8790–8797. doi:10.1128/JVI.71.11.8790-8797.1997

Winnerdy FR, Bakalar B, Maity A, Vandana JJ, Mechulam Y, Schmitt E, Phan AT. 2019. NMR solution and X-ray crystal structures of a DNA molecule containing both right- and left-handed parallel-stranded G-quadruplexes. *Nucleic Acids Res* **47:** 8272–8281. doi:10.1093/nar/gkz349

Yang SY, Lejault P, Chevrier S, Boidot R, Robertson AG, Wong JMY, Monchaud D. 2018. Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. *Nat Commun* **9:** 4730. doi:10.1038/s41467-018-07224-8

Yatime L, Maasch C, Hoehlig K, Klussmann S, Andersen GR, Vater A. 2015. Structural basis for the targeting of complement anaphylatoxin C5a using a mixed L-RNA/L-DNA aptamer. *Nat Commun* **6:** 6481. doi:10.1038/ncomms7481

Zhang N, Gorin A, Majumdar A, Kettani A, Chermichenko N, Skirpin E, Patel DJ. 2001. V-shaped scaffold: a new architectural motif identified in an A•(G•G•G•G) pentad-containing dimeric DNA quadruplex involving stacked G(anti)•G(anti)•G(anti)•G(syn) tetrads. *J Mol Biol* **311:** 1063–1079. doi:10.1006/jmbi.2001.4916

Zhou W. 2012. The new face of anaphylatoxins in immune regulation. *Immuobiology* **217:** 225–234. doi:10.1016/j.imbio.2011.07.016