



**HAL**  
open science

## Structure and functional reselection of the Mango-III fluorogenic RNA aptamer

Robert Trachman, Alexis Autour, Sunny Jeng, Amir Abdolazadeh, Alessio Andreoni, Razvan Cojocaru, Ramil Garipov, Elena Dolgosheina, Jay Knutson, Michael Ryckelynck, et al.

► **To cite this version:**

Robert Trachman, Alexis Autour, Sunny Jeng, Amir Abdolazadeh, Alessio Andreoni, et al.. Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nature Chemical Biology*, 2019, 15 (5), pp.472-479. 10.1038/s41589-019-0267-9 . hal-02112063

**HAL Id: hal-02112063**

**<https://hal.science/hal-02112063>**

Submitted on 28 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Published in final edited form as:

Nat Chem Biol. 2019 May ; 15(5): 472–479. doi:10.1038/s41589-019-0267-9.

## Structure and functional reselection of the Mango-III fluorogenic RNA aptamer

Robert J. Trachman III<sup>1</sup>, Alexis Autour<sup>2,†</sup>, Sunny C.Y. Jeng<sup>3,†</sup>, Amir Abdolazadeh<sup>3</sup>, Alessio Andreoni<sup>1</sup>, Razvan Cojocaru<sup>3</sup>, Ramil Garipov<sup>3</sup>, Elena V. Dolgosheina<sup>3</sup>, Jay R. Knutson<sup>1</sup>, Michael Ryckelynck<sup>2</sup>, Peter J. Unrau<sup>3</sup>, Adrian R. Ferré-D'Amaré<sup>1,\*</sup>

<sup>1</sup>Biochemistry and Biophysics Center, National Heart, Lung, and Blood Institute, 50 South Drive MSC 8012, Bethesda, MD 20892-8012, USA

<sup>2</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, Institut de biologie moléculaire et cellulaire du CNRS, 15 rue René Descartes, 67084, Strasbourg, France

<sup>3</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6

### Abstract

Several turn-on RNA aptamers that activate small molecule fluorophores have been selected *in vitro*. Among these, the ~30 nucleotide Mango-III is notable because it binds the thiazole orange derivative TO1-Biotin with high affinity and fluoresces brightly (quantum yield 0.55). Uniquely among related aptamers, Mango-III exhibits biphasic thermal melting, characteristic of molecules with tertiary structure. We report crystal structures of TO1-Biotin complexes of Mango-III, a structure-guided mutant Mango-III(A10U), and a functionally reselected mutant  $\Delta$ Mango-III. The structures reveal a globular architecture arising from an unprecedented pseudoknot-like connectivity between a G-quadruplex and an embedded non-canonical duplex. The fluorophore is restrained into a planar conformation by the G-quadruplex, a lone, long-range *trans*-Watson-Crick pair (whose A10U mutation increases quantum yield to 0.66), and a pyrimidine perpendicular to the nucleobase planes of those motifs. The improved  $\Delta$ Mango-III and Mango-III(A10U) fluoresce ~50% brighter than enhanced green fluorescent protein, making them suitable tags for live cell RNA visualization.

\*Address correspondence to A.R.F.; adrian.ferre@nih.gov.

<sup>†</sup>A. Au. and S.C.Y.J. contributed equally to this work

#### Author Contributions

R.J.T. performed all crystallographic studies, analytical ultracentrifugation, dynamic light scattering, fluorescence D<sub>2</sub>O titrations, and determined extinction coefficients and quantum yields. R.J.T. and S.C.Y.J. performed thermal melts. A.An. performed lifetime measurements and analysis. A.Ab., R.C., R.G., and S.C.Y.J. performed binding experiments. S.C.Y.J. performed SELEX, and A.Au. performed microfluidic reselection and bioinformatic sequence analysis. A.R.F. and R.J.T. wrote the manuscript with contribution from all authors (A.Au., S.C.Y.J., A.Ab., A.An., R.C., R.G., E.V.D., J.R.K., M.R., P.U.). All authors approved the final version of the manuscript.

#### Competing Financial Interests

The authors declare no competing financial interests.

## Introduction

Several RNA aptamers that strongly induce the fluorescence of their cognate small molecule fluorophores have recently been selected *in vitro* and used as molecular tags to image RNAs *in vivo*<sup>1-5</sup>. Such aptamers have the potential to transform the study of cellular RNAs, analogous to how fluorescent proteins revolutionized the study of the proteome<sup>6</sup>. For imaging applications, two key properties are the affinity with which the RNAs bind their cognate fluorophores, and the magnitude of fluorescence enhancement resulting from RNA binding. An *in vitro* selection experiment that explicitly sought<sup>3</sup> to isolate RNA aptamers that bind with high affinity to the thiazole orange (TO) derivative TO1-Biotin yielded 'RNA Mango'. This ~30 nucleotide (nt) aptamer binds TO1-Biotin, and the related TO3-Biotin (Supplementary Fig. 1) with high affinity (dissociation constant,  $K_d$ , of 3.0 and 8.0 nM, respectively). Crystallographic structure determination of the RNA Mango-TO1-Biotin complex revealed an RNA comprised of a three-tiered G-quadruplex linked to an A-form duplex through a flexible junction<sup>7</sup>. Consistent with the modest quantum yield (0.14) of this complex, the structure revealed that the methylquinoline (MQ) and benzothiazole (BzT) heterocycles of TO1-Biotin bind making a 45° angle relative to each other. Since a planar conformation should improve quantum yield<sup>8,9</sup>, it was hypothesized that RNA Mango variants may exist that induce brighter fluorescence from the same fluorophores<sup>10</sup>.

To discover brighter variants that also bind to TO1-Biotin with high affinity, the final pool from which RNA Mango (hereafter, Mango-I) was originally selected was subjected to a microfluidic, droplet-based functional selection, in which sequence variants were explicitly enriched for both brighter fluorescence and high affinity<sup>5</sup>. Three aptamers, Mango-II, Mango-III, and Mango-IV were isolated from these experiments and their properties characterized *in vitro* and *in vivo*. Of these, Mango-III is noteworthy for having the most intense fluorescence when bound to TO1-Biotin, being 40% brighter than enhanced green fluorescent protein (EGFP). This fluorescent tag may already be sufficiently bright for *in vivo* imaging techniques<sup>11</sup>.

To elucidate the molecular basis for the induction of bright fluorescence by Mango-III, and as a starting point for structure-guided optimization of this tool for live-cell imaging of RNAs, we now determined the co-crystal structure of Mango-III. Structure-guided mutation produced an 18% brighter mutant, Mango-III(A10U), and functional reselection produced  $\Delta$ Mango-III, a variant 13% brighter than Mango-III. Structures of these three RNAs revealed an unusually complex molecular architecture for a small RNA, including a pseudoknot-like base pairing interaction between nucleotides internal and adjacent to a two-tiered G-quadruplex. Mutational, photophysical, and reselection experiments demonstrated that nucleotides distant from the fluorophore were found to modulate properties of the aptamer-fluorophore complex.

## Results

### Overall Structure of Mango-III bound to TO1-Biotin

A 38-nt construct comprising the conserved 28-nt Mango-III core flanked by nucleotides presumed to form a 5-base pair Watson-Crick duplex was co-crystallized with TO1-Biotin.

The structure was solved by the single-wavelength anomalous dispersion (SAD) method using data from an iridium derivative (Supplementary Table 1, Methods). The crystallographic asymmetric unit (A.U.) contains two similar RNA-fluorophore complexes (RMSD = 0.15 Å for all non-hydrogen atoms). The two RNAs in the AU each make three crystal contacts involving end-to-end stacking of duplexes (Supplementary Fig. 2). Analytical ultracentrifugation and dynamic light scattering (Supplementary Fig. 3) however, indicate that Mango-III is monomeric in solution.

The structure of Mango-III (Fig. 1a,b) is organized around a two-tiered G-quadruplex with all-parallel connectivity, except for G18 of the top tier (T2), which is antiparallel. This quadruplex, which coordinates a canonical axial  $K^+$  ion ( $M_A$ , Supplementary Fig. 4), stacks on a base triple, which in turn stacks coaxially on an A-form duplex (paired element P1). Two nucleotides from the loop connecting G14 and G18 of T2, and four nucleotides from the loop that follows G20 (the last of the eight guanines of the G-quadruplex) form the three base pairs of a second paired element (P2) that is juxtaposed with the G-quadruplex. U17, which is extruded from the middle of P2, forms a *trans* Watson-Crick pair with A10 from the propeller loop that connects G9 to G13, on the diagonally opposite side of the T2 G-quartet. The BzT and MQ of the bound TO1-Biotin are sandwiched between the A10•U17 tertiary base pair and T2 of the G-quadruplex. Reflecting the high structural complexity of this RNA, only two nucleobases (A6, U11) of the 28-nt conserved core of Mango-III do not share at least a hydrogen bond with another nucleobase. In comparison, six nucleobases of the 22-nt core of Mango-I are fully unpaired (Fig. 1c).

### A coaxial duplex-triplex-quadruplex stack

Mango-I and the closely related Mango-II are each comprised<sup>7,12</sup> of a three-tiered G-quadruplex flexibly linked to an A-form helix through a GAA $\hat{A}$ -tetraloop-like junction (quadruplex insertion site,  $\hat{A}$ ). The residues of Mango-III corresponding to those junctions (Fig. 1c) instead fold into P1 by Watson-Crick base-pairing with complementary nucleotides at the 3' end of the aptamer (Fig. 1c). This is consistent with covariation analysis (ref. 5 and Supplementary Table 2). A25 of Mango-III (which corresponds in sequence to the last tetraloop residue of Mango-I) forms a base triple with A7 (which is part of the propeller loop between G5 and G8 of the Mango-III quadruplex) and U19 (which connects G18 and G20 of the Mango-III quadruplex) (Fig. 1a, 2a). The A7•U19•A25 triple is unusual because only two direct hydrogen bonds connect the three nucleobases, but each base also hydrogen bonds through its Watson-Crick face to a shared water molecule ( $W_1$ ) that lies near the center of the triple (Fig 2a).

The A7•U19•A25 base triple links the P1 duplex with the quadruplex of Mango-III through three sets of interactions. First, the base triple stacks below the T1 of the quadruplex and above the A3•U26 Watson-Crick base pair that closes P1 (Fig. 2a,b). Second, the 2'-OH of U26 donates a hydrogen bond to the N1 of A7 of the triple, in effect expanding it to a pseudo-tetrad. Third, a hydrated cation ( $M_B$ , modeled as  $Na^+$  due to lack of anomalous signal with 1.495 Å X-rays and coordination<sup>13</sup> by N3 of a purine) and a network of ordered water molecules connects the phosphate of A7 to the minor groove faces of both, A3 and

G4, thereby spanning the duplex-quadruplex junction (Fig. 2c). These interactions likely produce a rigid connection between P1 and the G-quadruplex moieties of Mango-III.

### A partially parallel non-canonical duplex

The P2 duplex of Mango-III is unique because of the predominantly parallel orientation of its strands, and because its three base pairs are non-canonical (Fig. 2d,e,f). The backbones of the *trans* Watson-Crick U15•U22 and *trans* Hoogsteen A16•A23 base pairs are parallel. The third base pair is formed between U24 and G21, whose backbones are locally anti-parallel, being on opposite ends of the same loop connecting G20 of T1 (the last guanine of the G-quadruplex) to A25 of the base triple. The G21•U24 pair is formed in *cis*, and comprises a single hydrogen bond. The unusual structure of the P2 duplex is reflected in C1'-C1' distances (8.5 Å, 11.4 Å and 12.6 Å, respectively, for the three pairs), different from that of canonical antiparallel A-form RNA duplexes (10.7 Å), and appears to be stabilized by a cross-strand stacking of A16 on G21. This latter interaction gives rise to reciprocal hydrogen bonds between N1 of A16 and 2'-OH of G21, and O6 of G21 and 2'-OH of A16.

The P2 helix contributes to the overall structure of Mango-III and the precise conformation of its fluorophore binding pocket in three ways. First, because P2 is followed immediately by A25, the duplex anchors the two-tiered G-quadruplex onto the base triple. Second, G18 and G20 form the only non-contiguous and anti-parallel guanine stack in the Mango-III G-quadruplex, and P2 anchors the backbones of these two nucleotides next to each other. This may be further facilitated by a hydrogen bond between the *pro-R<sub>p</sub>* non-bridging phosphate oxygen (NBPO) of U15 and the 2'-OH of G14 from the adjacent guanine stack (Fig. 3a). Third, P2 helps position U17 of the tertiary base pair in the fluorophore binding pocket of Mango-III. Mutational analysis of P2 indicates that positions 16, 21, 23, 24 are sensitive to base identity while position 15 is less so (Supplementary Table 3).

### The TO1-Biotin-binding site of Mango-III

Unbiased residual electron density corresponding to the BzT and MQ, as well as the four proximal atoms of the PEG linker of TO1-Biotin unambiguously located the TO headgroup between T2 of the G-quadruplex and the tertiary A10•U17 *trans* Watson-Crick base pair. A strong anomalous difference Fourier feature corroborated the location of the sulfur of BzT (Fig. 3b). In its Mango-I complex, the planes of BzT and MQ of the fluorophore subtend a 45° angle, and its biotin moiety packs against MQ and the G-quadruplex of the aptamer. In contrast, when bound to Mango-III, the BzT and MQ rings are coplanar in both copies of the AU (Fig. 3b), and the biotin projects out into solvent, and lacks electron density.

To test the functional significance of the tertiary A10•U17 base pair of Mango-III, we generated the fifteen additional combinations of natural nucleotides at these two positions, as well as substitutions of A10 with 2-aminopurine and 2,6-diaminopurine, and examined their effect on TO1-Biotin binding and fluorescence enhancement (Supplementary Table 4). With one exception, these changes resulted in >10-fold decrease in binding affinity and >24% loss in fluorescence enhancement. The exception was A10U, which not only exhibited TO1-Biotin binding affinity indistinguishable from that of the wild-type, but also a 18% higher fluorescence enhancement. Thus, the tertiary base pair of Mango-III is important for

fluorescence enhancement, of the sixteen possible nucleotide pairs, U10•U17 yields the brightest fluorescence.

Binding of TO1-Biotin to Mango-III appears to be further stabilized by U12. Its nucleobase is perpendicular to the planes of T2 and the A10•U17 pair, and it hydrogen bonds through its N3 to the N3 of A10, and through its 2'-OH to the *pro-S<sub>p</sub>* NBPO of G14 (Fig. 3a,b). The nucleobase of U12 is also in van der Waals contact with the proximal atoms of the PEG linker of TO1-Biotin. Previous characterization of Mango-III showed that the U12C mutant was 40% less fluorescent<sup>5</sup>. We prepared U12A and abasic residue 12 variants and found that they exhibited >30-fold loss of affinity and >50% loss in fluorescence enhancement (Supplementary Table 4). Thus, residue 12 is important for Mango-III function, and a uracil at this position appears to be optimal for TO1-Biotin activation (the A10U, U17A double mutant of Mango-III improves fluorescence enhancement of TO3-Biotin by 30%, Supplementary Table 4).

### Structure of the enhanced Mango-III(A10U) mutant

To determine the structural basis of the enhanced fluorescence of the A10U mutant, we solved its crystal structure in complex with TO1-Biotin at 2.9Å resolution (Supplementary Table 1, Methods, and Fig 3c, d). This mutant crystallized with four independent aptamer-fluorophore complexes in the A.U. The four complexes have similar overall structures; nonetheless, two of them (chains B and D) have P2 helices in which the cross-strand-stacked A16 and G21 are ordered, but U15, A23 and U24 are substantially disordered (Supplementary Fig. 5). In all four chains, the U10 base resulting from the mutation engages in a *trans* Watson-Crick base pair with U17. The C1'-C1' distance of the mutant U10•U17 tertiary base pair (11.1 Å) is largely unchanged from that of the parental A10•U17 base pair (11.2 Å). However, the smaller pyrimidine at position 10 of the mutant reduces solvent occlusion of MQ of the fluorophore while retaining the H-bonding to U12. This and the partial disorder of P2 are reflected in a reduced shape complementarity<sup>14</sup> between the A10U mutant and the fluorophore headgroup, compared to wild-type (S.C. scores of  $0.837 \pm 0.03$  and  $0.854 \pm 0.01$ , respectively; mean  $\pm$  r.m.s.d.) Consistent with the looser packing, the TO headgroups of the bound TO1-Biotin fluorophores adopt the (*E*) conformation, as seen in Mango-III, in two of the complexes in the A10U mutant crystal structure, while the other two are in the (*Z*) conformation (Fig. 3d).

### Structure-guided reselection of Mango-III

The improved fluorescence enhancement of Mango-III(A10U) led us to hypothesize that other improved variants of Mango-III may exist. To test this, we subjected the aptamer to structure-guided functional reselection (Fig. 4). A starting pool of  $4.1 \times 10^6$  sequences was generated by mixing two RNA pools in equal proportion, with pool 1 randomized at ten positions of the conserved Mango-III core (residues 3, 7, 10, 12, 17, 19, 21, 24, 25, 26), and pool 2 randomized at the same positions as pool 1 with an additional random insertion 3' to position 26. We reasoned the latter could support a variant base quadruple (Fig. 4a). Two rounds of SELEX to enrich for high-affinity aptamers<sup>15,16</sup> were followed by four rounds of droplet-based screening<sup>17</sup> for bright aptamers. The selection converged on four dominant sequences with frequencies of 58%, 12%, 8.5%, and 1.2%. The four dominant reselected

Mango-III variants were tested for TO1-Biotin binding and fluorescence enhancement. They exhibited dissociation constants of 3.6 – 6.4 nM and fluorescence enhancements (relative to wild-type) of –2% to +15% (Supplementary Table 5, 6). The most populated variant in the selection (*Δ*Mango-III) exhibited fluorescence enhancement 13% higher than that of wild-type.

All four functionally reselected sequences share uridines at positions 10 and 17, indicating that the tertiary base pair of Mango-III(A10U) is already optimal. All reselected sequences retain the insertion after residue 26 (hereafter residue 26:1) and in all cases this insertion is a guanine. The reselected sequences also differ from wild-type and Mango-III(A10U) in their duplex-triplex-quadruplex junction and in P2. Residue 7, which as an adenine participates in the junctional base triple in the wild-type, reselected as either adenine or cytosine, residue 19 (as a uracil, part of the wild-type triple) reselected as guanine, uracil or adenine, and residue 25 (which as an adenine completes the wild-type triple), reselected as an uracil. All four sequences harbor a guanine at position 24, which in the wild-type is occupied by a uracil that makes the G21•U24 closing base pair of P2. Collectively, these changes suggest that the structure of the junction region of the aptamer was altered.

### Structure of the functionally reselected *Δ*Mango-III

To determine the molecular consequences of functional reselection, we determined the structure of *Δ*Mango-III in complex with TO1-Biotin at 1.55 Å resolution (Table 1, Online Methods). The structure reveals three sets of interrelated changes that propagate from G26:1, which forms a *cis* Watson-Crick pair with A3, thereby closing the G-quadruplex-proximal end of P1. U26, rather than pairing with A3 as in Mango-III and Mango-III(A10U), forms a base triple comprised of the A7•U26 Watson-Crick pair, and G19 (Fig. 5a). The four reselected variants of Mango-III can all form this triple (Supplementary Fig. 6). The repositioning of residue 26 into the triple, made possible by the presence of G26:1, extrudes residue 25 from the duplex-triplex-quadruplex junction (in wild-type, A25 is part of the junctional base triple) (Fig. 5b). Residue 25, a uridine in *Δ*Mango-III, stacks below P2, and anchors it to the quadruplex by making three hydrogen bonds with G13 of T1 while the residue at position 24, a guanine in *Δ*Mango-III, forms a *cis* Hoogsteen base pair with G21, which in turn stacks below the *trans* Watson-Crick A16•A23 pair (Fig 5c). As in the wild-type Mango-III, A16 and G21 make a cross-strand stack in P2 of *Δ*Mango-III, but in the latter RNA, the two homopurine pairs (G21•G24 and A16•A23) are severely propeller-twisted compared to the near planar pairs in P2 of the wild-type. This twisting and a 28° rearrangement of the P2 and G-quadruplex helical axes (Supplementary Fig. 7), possibly abetted by the anchoring of the bottom of P2 by U25, allows G21 and G24 to make three direct and three water-mediated hydrogen bonds with the adjacent groove of T2 of the G-quadruplex (Fig 5d). None of these interactions are present in wild-type. P2 of *Δ*Mango-III lacks the closing U15•U22 base pair. Instead, U22 is extruded into solvent and its nucleobase disordered, and U15 packs against the distal portion of the PEG linker of the bound TO1-Biotin, which in this crystal structure has defined electron density nearly in its entirety (Fig 5e).

We tested the functional significance of  $\Delta$ Mango-III structural features by mutating P2 and the junctional triple. All P2 mutations reduced binding affinities to TO1-Biotin ( $K_d$ 's from 32 nM to 338 nM) and fluorescence enhancements (by 12% – 40%). Mutations to the junctional triple that were expected to maintain stability all had a marginal loss of fluorescence enhancement and affinity for TO1-Biotin, while mutations expected to disrupt the base triple resulted in  $K_d$ 's ranging from 137 nM – 408 nM, and reduced fluorescence enhancement by up to 20% (Supplementary Table 7).

### Photophysical properties of Mango-III and variants

Despite the 18% and 13% increase in fluorescence enhancement by Mango-III(A10U) and  $\Delta$ Mango-III, respectively, relative to wild-type, as well as more than 20% of the core nucleotides being different in the functionally re-selected aptamer, Mango-III and its two variants exhibit indistinguishable affinities for TO1-Biotin (Supplementary Tables 4, 5). Consistent with this, there is no difference in thermal stability between them, and all three have the same profile for the first (tertiary) melting transition (Supplementary Fig. 8). The absorbance and fluorescence maxima of  $\Delta$ Mango-III are blue-shifted by 4 nm and 7 nm, respectively (Fig. 6a,b). To test whether the blue-shift of  $\Delta$ Mango-III is dependent on the U•U tertiary base pair, we generated an  $\Delta$ Mango-III(U10A) mutant in which the A•U tertiary base pair of the wild-type is restored. This mutant retains the spectrum of  $\Delta$ Mango-III while the spectrum of Mango-III(A10U) is identical to that of Mango-III (Supplementary Fig. 9), indicating that the  $\Delta$ Mango-III blue-shift results from its reselected core. Further photophysical analysis shows that the three RNA-fluorophore complexes have identical extinction coefficients, and the improved fluorescence enhancement by Mango-III(A10U) and  $\Delta$ Mango-III results from higher quantum yields [0.55, 0.66 and 0.64 for the TO1-Biotin complexes of Mango-III, Mango-III(A10U) and  $\Delta$ Mango-III, respectively; Supplementary Table 6].

In principle, higher quantum yield should be accompanied by longer fluorescence lifetime<sup>18</sup>. This, however, is not the case for Mango-III and variants. While the wild-type exhibits a lifetime of ~4.9 ns, Mango-III(A10U),  $\Delta$ Mango-III, and  $\Delta$ Mango-III(U10A) all had equal or shorter lifetimes (Supplementary Table 8). The lack of proportionality between fluorescence lifetimes and quantum yields could be due to contributions from ultra-fast decays (<100 ps)<sup>19–21</sup> undetectable due to instrumental limitations (Online Methods). In an attempt to slow down potential molecular movements of the fluorophore, we performed lifetime measurements in solutions with increased viscosity. Reducing molecular motions has the potential to lessen ultra-fast decays, and in this case, help rationalize the discrepancy between quantum yields and lifetimes. Addition of either sucrose or glycerol, however, resulted in shorter average lifetimes and reduced brightness (Supplementary Fig. 10), suggesting that the structures of the RNAs are negatively perturbed by these solutes.

D<sub>2</sub>O can affect fluorescence lifetimes through a kinetic isotope effect on transfer of protons (deuterons)<sup>22</sup> or through changes in viscosity. Since the protons of TO1-Biotin do not readily exchange, effects of D<sub>2</sub>O should arise from the latter under constant pH (or pD). To detect additional lifetime decays through increased viscosity, we performed D<sub>2</sub>O titrations on the Mango-III-TO1-Biotin complexes. To control for differences in direct solvent

interaction with the fluorophore, we performed D<sub>2</sub>O titrations for  $\Delta$ Spinach and Corn, two aptamers with demonstrated fluorophore-H<sub>2</sub>O hydrogen bonding<sup>23,24</sup>, and the DIR-Pro aptamer, a promiscuous aptamer with demonstrated fluorescence enhancement of TO1-Biotin<sup>25,26</sup>. Titration with D<sub>2</sub>O resulted in a linear increase in fluorescence intensity for Mango-III and  $\Delta$ Mango-III (Fig. 6c) and the control aptamers, Corn,  $\Delta$ Spinach, and DIR-Pro (Supplementary Fig. 11). At 99.6% D<sub>2</sub>O, the fluorescence lifetimes of Mango-III and Mango-III(A10U) increased by 5.3% and those of  $\Delta$ Mango-III and  $\Delta$ Mango-III(U10A) by 9.2% (Fig. 6d, Supplementary Fig. 12), but additional lifetimes could not be resolved (Supplementary Table 8, Supplementary Fig. 13). The control aptamers exhibited longer lifetimes in D<sub>2</sub>O with Corn (14.0%)>DIR-Pro (6.1%)> $\Delta$ Spinach (1.3%) (Supplementary Table 9). Thus, no pattern was apparent between fluorophore-solvent hydrogen bonding and lifetimes.

## Discussion

Previous Mango-I and Mango-II co-crystal structures revealed a shared molecular architecture comprised of a G-quadruplex and an A-form duplex (P1) flexibly linked through a tetraloop-like junction<sup>7,12</sup>. Despite its comparable size, the 3D structure of Mango-III is considerably more elaborate, incorporating a second helical element (P2), a long-range *trans*-Watson-Crick base pair, as well as a coaxially stacked junction between P1 and the G-quadruplex (Fig. 1). The formation of P2 by non-canonical base pairing between two nucleotides in a lateral loop of the G-quadruplex and four nucleotides immediately 3' to it gives rise to a molecular connectivity that is analogous to that of a pseudoknot. A canonical pseudoknot arises from Watson-Crick base pairing between nucleotides in the loop of a stem-loop and those in an adjacent single-stranded segment<sup>27</sup>. The pseudoknot-like G-quadruplex fold of Mango-III is unprecedented. The complex tertiary structure of Mango-III explains why this aptamer, unlike Mango-I or Mango-II, exhibits a biphasic thermal melting profile<sup>5</sup>. As with proteins<sup>28</sup>, the requirement of structural complexity<sup>29,30</sup> sufficient to make multidirectional interactions with the fluorophore in order to achieve bright fluorescence may set a lower bound on the size of fluorogenic RNA aptamers. At ~30 nt, almost all core nucleobases of Mango-III and  $\Delta$ Mango-III engage in hydrogen bonding. Perhaps not coincidentally, 30 nt is the lower limit in size of riboswitch aptamer domains<sup>31–33</sup> and may be close to the minimal size for an RNA to possess tertiary structure and selectively bind a small molecule.

Most of the structurally characterized fluorescence turn-on aptamers incorporate G-quadruplex motifs within their functional cores (exceptions being<sup>34,25,35</sup> malachite green, DIR, and Riboglow). G-quadruplexes can provide a large surface for multiple heterocycles to bind in-plane, thus maximizing fluorescence enhancement. However, *in vivo*, G-quadruplexes may suffer from off-target interactions with helicases<sup>36</sup>. Recent structural characterization<sup>37</sup> of the substrate-bound mammalian helicase DHX36 demonstrated that it recognizes both, the guanine bases and the phosphodiester backbone of a simple, all-parallel G-quadruplex. By comprising flanking A-form P1 and irregular P2 helices, as well as a *trans* Watson-Crick base pair around a mixed-connectivity G-quadruplex core, Mango-III may exhibit reduced interactions with cellular G-quadruplex-binding proteins that may destabilize simpler G-quadruplex-containing RNAs.