DOI: 10.1002/cbic.200800183

TokenRNA: A New Type of Sequence-Specific, Label-Free Fluorescent Biosensor for Folded RNA Molecules

Kirill A. Afonin, Evgeny O. Danilov, Irina V. Novikova, and Neocles B. Leontis*[a]

This paper is dedicated to Marina V. Afonin.

Fluorescent reporters are highly sensitive, nonperturbing, and convenient probes for biological studies. Current techniques for recognition of specific nucleic acid sequences usually require complementary hybridization to chemically modified probes.[1] Optimal fluorescent biosensors should rapidly signal the presence of a specific analyte with high selectivity and high contrast. Recently, a number of label-free, fluorescent bioindicators have been reported including aptamer-based riboswitches as sensors for cofactors,[2] aptamer-based protein sensors,[3] and chimeric aptamers, in which the recognition domain binding the target was coupled to an aptamer that binds a fluorophore in such a way that the presence of the target increases the affinity of the bioindicator for the fluorophore.[4,5] When the emission yields of the fluorophore in its bound and free states differ, this provides a binary sensor with nonperturbing optical recognition.

It is desirable to detect macromolecular analytes (for example, structured RNA molecules) in their native environment without having to denature or unfold them. In this work, we demonstrate a technique for sensitive, label-free, real-time sequence-specific recognition of prefolded RNA sequences. We use RNA constructs designed to form an aptameric pocket for the fluorophore upon programmable paranemic binding to a specific prefolded analyte RNA sequence. We call these constructs paranemic "token RNAs".

We chose the triphenylmethane dye, Malachite Green (MG), as the fluorescent reporter because in its unbound state in water solution it exhibits extremely low fluorescence quantum yield from the S1 excited state because of efficient internal conversion.[6,7] The emission of the dye increases substantially when the nonradiative relaxation channels from S1 are shut down. Whereas the detailed underlying mechanisms of this phenomenon are still being debated,[7–9] related studies show that "rigidifying" the dye by placing it in a highly viscous environment or in a binding cage increases its emission dramatically.[7] For instance, it was reported recently that the emission of MG increases by several orders of magnitude upon binding to an RNA aptamer obtained by in vitro selection (SELEX).[8,10] This aptamer has also been used as a reporter for ATP recognition[4] and to perform real-time fluorescent monitoring of single-stranded DNA molecules.[11] The DNA detection was based on the separation of the MG aptamer into two strands, each of which is linked to a nucleic acid arm complementary to one half of the analyte DNA sequence so as to form a three-way junction in the presence of the target. In the presence of the analyte, the intensity of MG fluorescence substantially increased. This probe was sensitive to a single nucleotide substitution.[11]

The utility of DNA 5HT and 7HT paranemic cohesions for assembling 2D DNA arrays is well established.[12–14] We designed our sensor tokenRNA by coupling the malachite green or "MG" aptamer to the three-half-turn (3HT) RNA paranemic binding motif we recently designed and characterized.[15] The 3D structure of the MG-aptamer (PDB files: 1f1t and 1q8n) shows that the MG binding pocket comprises an internal loop embedded within the RNA helix.[16] The 3D structure of the aptamer is represented in Figure 1 using the annotations for non-Watson–Crick basepairs proposed by Leontis and Westhof.[18] The numbering of the nucleotides of the aptamer in Figure 1 is consistent with their numbering in the biosensor, **36Mtoken** (see Scheme 1).

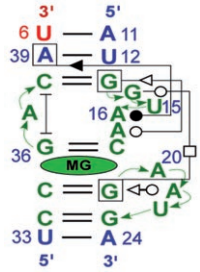


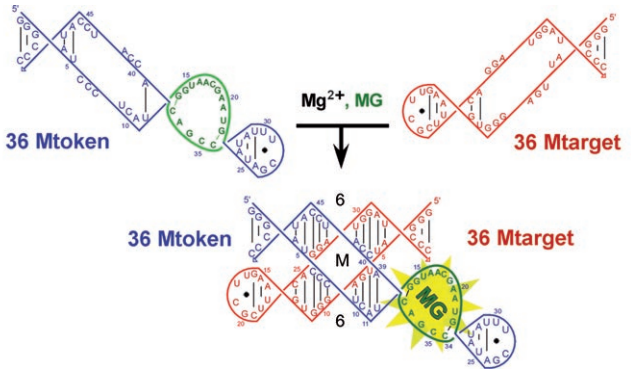
Figure 1. Annotation of the MG aptamer 3D structure using the Leontis-Westhof basepair nomenclature.[18] Non-Watson–Crick basepairs are annotated to show interacting edges: Watson–Crick (WC) Edge (circles), Hoogsteen Edge (squares), or the Sugar Edge (triangles). Open symbols indicate *trans* pairs and solid symbols, *cis* pairs. Boxes are placed around bases involved in more than one basepair interaction.[19] The bases of the aptamer are shown in green.

As shown in Figure 1, bound MG intercalates between a Watson–Crick (WC) basepair (C18/G36) and a base quadruple (G14/G19A21/C35), and makes additional contacts with unpaired bases. For example, unpaired A20 stacks on the MG phenyl group and interferes with its rotation.

Using computer 3D modeling, we positioned the aptamer internal loop close to the paranemic binding site and oriented it so the Watson–Crick helix flanking the aptamer and stabilizing it so that it can only form when the sensor RNA binds to a target (analyte) RNA that contains the complementary sequence to form the paranemic motif. When the sensor RNA binds the target RNA, the intermolecular basepairs of the paranemic motif form adjacent to the aptamer, starting with U6 (target) pairing with A11 (sensor). These basepairs stabilize the A39/U12 basepair of the tokenRNA, which interacts with A16 in the aptamer motif to form the A16A39 *cis* Watson–Crick/Sugar Edge (cWS) tertiary basepair. Other nearby tertiary inter-

[a] *K. A. Afonin, E. O. Danilov, I. V. Novikova, Prof. N. B. Leontis*
Department of Chemistry ,Bowling Green State University
Bowling Green, OH 43403 (USA)
Fax: (+1) 419-372-9809
E-mail: leontis@bgsu.edu

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.



Scheme 1. Schematic representation of the assembly of **36Mtoken** and **36Mtarget** to form the recognition complex, stabilizing the Malachite Green (MG) aptamer site (see Figure 1).

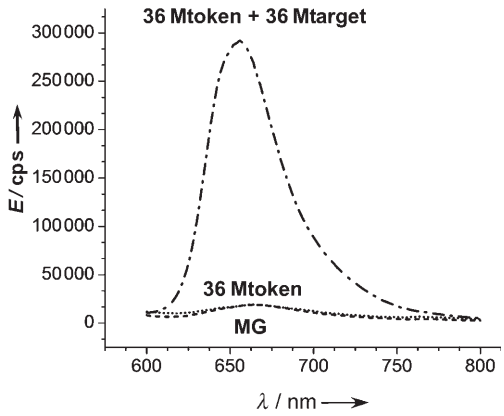


Figure 2. MG aptamer enhances the fluorescence of MG upon the assembly of **36Mtoken/36Mtarget** recognition complex. Emission spectra of free MG (-----), MG in the presence of aptamer-containing **36Mtoken** only (····), and MG in the presence of the recognition complex **36Mtoken/36Mtarget** containing the stabilized MG aptamer site (–·–·–).

actions that stabilize the active form of the aptamer includes A17/G13 (tWS), G14/G19 (tWH), and A21/G19 (tWS). Scheme 1 shows how the biosensor RNA (**36Mtoken**) binds to the target (**36Mtarget**) to form the paranemic 3HT recognition complex.

RNA paranemic molecules are designated according to the number of half-turns in the paranemic motif (three in this case) and the number of basepairs separating the strand crossover points across the major (M) groove (six in this case).[12,15] "Token" in the name of a molecule indicates the sensor RNA in which a reporter module is placed (the MG aptamer in this case). "Target" indicates the RNA molecule recognized by the tokenRNA. For example, **36Mtoken** means that this molecule forms a 3HT complex with six basepairs spanning the major (M) groove.

All RNA molecules were run-off transcribed using T7 RNA polymerase and PCR-amplified DNA templates having appropriate promoters, and subsequently purified on denaturing gels. Native gel electrophoreses was used as previously described[20] to assay RNA assembly in the presence and absence of MG. The previously reported paranemic dimer **36**M**1**/**36M2** was used as a gel mobility control.[15] (**36M2** is identical to **36Mtarget** in this report whereas **36Mtoken** is **36M1** plus the MG aptamer.) These experiments showed that **36Mtarget** binds to cognate **36M1** or **36Mtoken** in the presence or absence of MG (see the gel presented in Figure S1 in the Supporting Information).

Next we measured MG emission in the presence and absence of **36Mtarget** and **36Mtoken**. MG was excited at 425 nm and fluorescence measured using a Jobin Yvon Fluorolog 3-11 spectrofluorimeter, as detailed in the Experimental Section.

As shown in Figure 2, when MG was added to the aptamer-containing tokenRNA, **36Mtoken** in the absence of the RNA target **36Mtarget**, the MG emission remained very low, similar to that of free MG. However, when MG was added to **36Mtoken** in the presence of **36Mtarget**, the fluorescence of MG increased dramatically. This indicates that paranemic assembly of the sensor with the target is required to stabilize the MG aptamer pocket. The stabilization probably occurs through a cascade of tertiary interactions triggered by the formation of the intermolecular U6A11 basepair (see above).

To determine the specificity of the sensor for its target, point mutations were introduced in the target molecule **36Mtarget** (Figure 3 and Scheme S1). The altered molecules were named **36Mtarget(a)**, which contains mutation G11→C, and **36Mtarget(c)**, which contains mutation U6→A. Molecules **36Mtoken(a)** and **36Mtoken(c)** were designed to pair with **36Mtarget(a)** and **36Mtarget(c)**, respectively. **36Mtoken(a)** contains mutation C6→G and **36Mtoken(c)** contains mutation A11→U. These mutations were placed in the major crossover grooves far from the aptamer motif so as not to directly affect its folding. In previous work we showed that mutations in these locations do not perturb the free energy of folding of the individual strands.[15]

In the presence of only **36Mtoken(a)** or **36Mtoken(c)**, MG does not emit (see Figure 3 for the **(a)** constructs and Figure S2 for the **(c)** molecules). Also, no increase in fluorescence of MG was observed in the presence of **36Mtoken** mixed with **36Mtarget(a)** or **36Mtarget(c)**. These results show that a single base mismatch in the complementary strand abolishes the MG emission signal. However, binding of the cognate token RNA that restored WC basepairing in the paranemic motif [**36Mtoken(a)/36Mtarget(a)** and **36Mtoken(c)/36Mtarget(c)**] also restored MG emission. These results indicate that the paranemic aptamer system is capable of single-nucleotide discrimination, in agreement with previous work in which we characterized the specificity of major groove RNA 3HT paranemic assembly.[15]

To show that we can systematically create new biosensors for prefolded RNA sequences utilizing RNA–RNA paranemic assembly and MG aptameric pocket stabilization, we designed two more target-RNAs [**36Mtarget(1)** and **36Mtarget(2)**] and two complementary token-RNAs [**36Mtoken(1)** and **36Mtoken(2)**]. The sequences of these constructs are shown in Scheme S1. In the presence of either **36Mtoken(1)** or **36Mtoken(2)** alone, the MG emission was insignificant but was found to increase substantially upon binding to the complementary token–targetRNA complex (Figures 3 and S4). More-

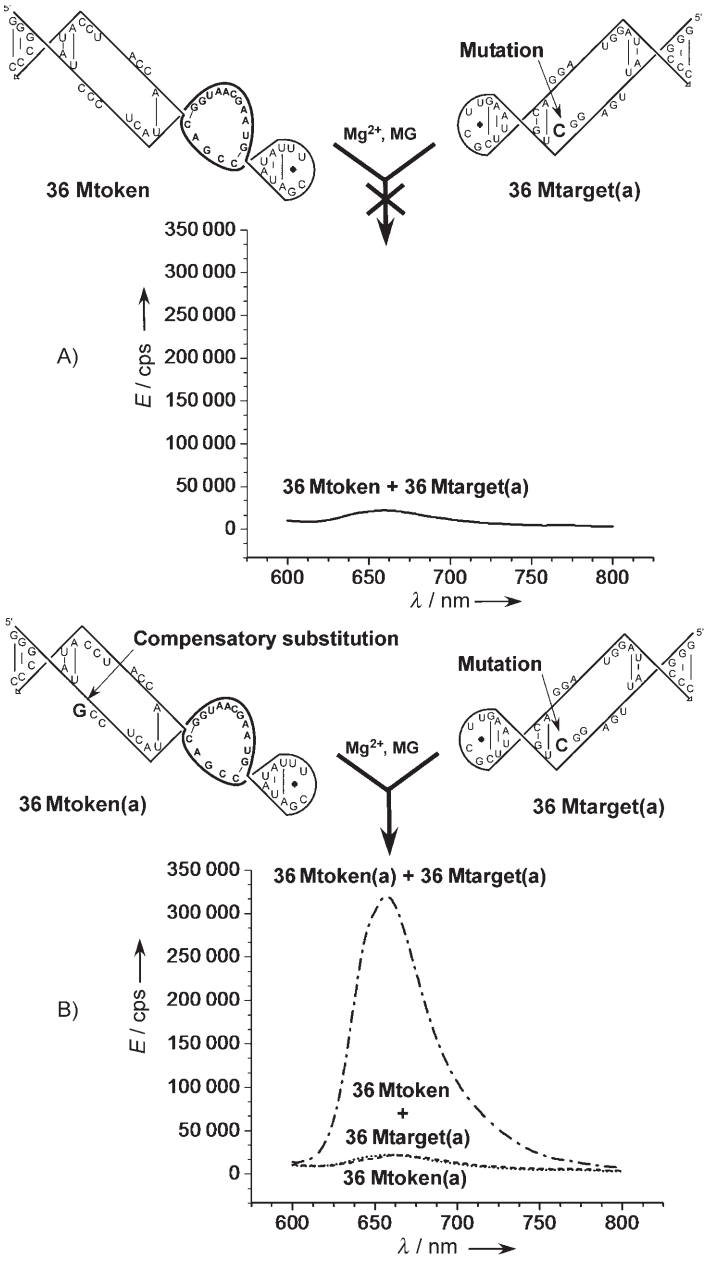


Figure 3. Schematic representation of the sequence specificity of target RNA recognition and supporting emission spectra. A) Single point mutation (G11→C) in the target, **36Mtarget(a)**, blocks recognition by **36Mtoken**. B) Compensatory substitution (C6→G) in **36Mtoken(a)** supports recognition of **36Mtarget(a)** and turns on Malachite Green (MG) emission signal.

over, mixing of **36Mtoken(1)** with noncognate **36Mtarget** or **36Mtarget(2)** had no effect on the MG emission signal (Figure S5). These results confirm our initial findings and prove the potential to design tailored tokenRNAs for specific targets.

In summary, we have shown how to systematically create new programmable biosensors targeting prefolded RNA sequences by programmable recognition and analysis using paranemic binding motifs coupled to the MG aptamer. We have demonstrated the high specificity of this system, which can discriminate among RNAs that differ from the target RNA by one nucleotide in the paranemic recognition domain. Finally, we have demonstrated the generality of our approach by designing tokenRNA biosensors for three completely different RNA targets.

Experimental Section

RNA preparation: RNA molecules were prepared by run-off transcription of PCR amplified DNA templates as previously described.[20] Synthetic DNA molecules coding for the antisense sequence of the desired RNA were purchased from IDT DNA (http://www.idtdna.com) and amplified by PCR using primers containing the T7 RNA polymerase promoter. PCR products were purified using the QiaQuick PCR purification kit (Qiagen Sciences, Maryland 20874). RNA molecules were prepared by in vitro transcription using T7 RNA polymerase (Takara Bio Inc. http://www.takara-bio.com) and purified on denaturing polyacrylamide gels (PAGE) (15% acrylamide, 8 M urea). The RNA was eluted from gel slices overnight at 4 °C into buffer (300 mM NaCl, 10 mM Tris pH 7.5, 0.5 mM EDTA), then ethanol precipitated, rinsed twice with ethanol (80%), dried, and dissolved in water.

Radiolabeling of RNA molecules: T4 phosphokinase (T4PK from New England BioLabs Inc.) was used to transfer the ^{32}P-gamma phosphate of ATP to the 5′-end of 3′-cytidine monophosphate (Cp) to form radiolabeled pCp. T4 RNA ligase (New England BioLabs Inc.) was used to attach radiolabeled pCp to the 3′-ends of RNA molecules (10–20 pmol). Labeled material was purified on denaturing polyacylamide gels (12% acrylamide, 19:1 bis:monomer, 8 M urea).

Assembly experiments: Prior to the addition of the buffer and $Mg(OAc)_2$, RNA samples containing a fixed amount (~0.5 nM) of ^{32}P-3′-end labeled RNA and the unlabeled binding partner molecule (200 nM) were heated to 90 °C for 1 min and snap cooled on ice to avoid intermolecular base pairing. Tris-borate buffer (89 mM, pH 8.3) was added and the samples were incubated at 30 °C for 5 min. Then $Mg(OAc)_2$ (15 mM) and MG (5 μM) were added and incubation continued for 30 min. An equal volume of loading buffer (same buffer with 0.01% bromophenol blue, 0.01% xylene cyanol, 50% glycerol) was added to each sample for analysis on 7% (15:1) polyacrylamide native gels (containing 15 mM $Mg(OAc)_2$) and run at 4 °C with constant recycling of the running buffer (89 mM Tris-borate, pH 8.3/15 mM $Mg(OAc)_2$). Gels were run for 3 h, at 50 mA with constant buffer recirculation, dried under vacuum, placed on a phosphor storage screen for 16 h, and scanned using a Storm phosphoimager (Amersham, Storm 860, http://www.gehealthcare.com).

Fluorescent experiments. The fluorescent experiments were carried out using a Jobin Yvon Fluorolog 3–11 spectrofluorimeter with the following settings: The excitation wavelength was set at 425 nm in all experiments. Emission was scanned from 600 to 800 nm in increments of 2 nm. Integration time was 1 s and detector voltage was 950 V. Signal was registered in counts per second (cps). Excitation and emission slits were both set to 4 nm.

Acknowledgements

The authors thank Professor Alexander N. Tarnovsky for helpful discussions and Professor Douglas C. Neckers for access to fluorescence instrumentation. This work was supported by a grant from the National Institutes of Health (2 R15M055898-04).

Keywords: aptamers · biosensors · fluorescence · paranemic · RNA recognition · tokenRNA

[1] R. M. Long, D. J. Elliott, F. Stutz, M. Rosbash, R. H. Singer, *RNA* **1995**, *1*, 1071.
[2] A. Ogawa, M. Maeda, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3156.
[3] Y. Jiang, X. Fang, C. Bai, *Anal. Chem.* **2004**, *76*, 5230.
[4] M. N. Stojanovic, D. M. Kolpashchikov, *J. Am. Chem. Soc.* **2004**, *126*, 9266.
[5] L. Wu, J. F. Curran, *Nucleic Acids Res.* **1999**, *27*, 1512.
[6] S. B. Baptista, G. L. Indig, *J. Phys. Chem. B* **1998**, *102*, 4678.
[7] D. F. Duxbury, *Chem. Rev.* **1993**, *93*, 381.
[8] J. R. Babendure, S. R. Adams, R. Y. Tsien, *J. Am. Chem. Soc.* **2003**, *125*, 14716.
[9] D. H. Nguyen, S. C. DeFina, W. H. Fink, T. Dieckmann, *J. Am. Chem. Soc.* **2002**, *124*, 15081.
[10] D. Grate, C. Wilson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6131.
[11] D. M. Kolpashchikov, *J. Am. Chem. Soc.* **2005**, *127*, 12442.
[12] X. Zhang, H. Yan, Z. Shen, N. C. Seeman, *J. Am. Chem. Soc.* **2002**, *124*, 12940.
[13] P. E. Constantinou, T. Wang, J. Kopatsch, L. B. Israel, X. Zhang, B. Ding, W. B. Sherman, X. Wang, J. Zheng, R. Sha, N. C. Seeman, *Org. Biomol. Chem.* **2006**, *4*, 3414.
[14] Z. Shen, H. Yan, T. Wang, N. C. Seeman, *J. Am. Chem. Soc.* **2004**, *126*, 1666.
[15] K. A. Afonin, D. J. Cieply, N. B. Leontis, *J. Am. Chem. Soc.* **2008**, *130*, 93–102.
[16] J. Flinders, S. C. DeFina, D. M. Brackett, C. Baugh, C. Wilson, T. Dieckmann, *ChemBioChem* **2004**, *5*, 62.
[17] C. Baugh, D. Grate, C. Wilson, *J. Mol. Biol.* **2000**, *301*, 117.
[18] N. B. Leontis, E. Westhof, *RNA* **2001**, *7*, 499.
[19] N. B. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res.* **2002**, *30*, 3497.
[20] K. A. Afonin, N. B. Leontis, *J. Am. Chem. Soc.* **2006**, *128*, 16131.

Received: March 25, 2008
Published online on July 24, 2008

