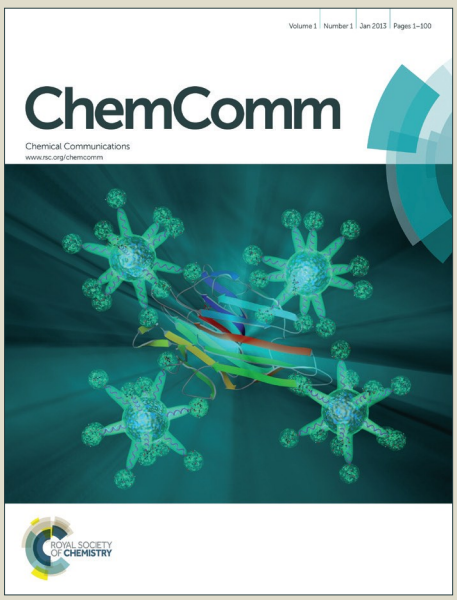
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Evolution and Characterization of a Benzylguanine-binding RNA Aptamer

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

J. Xu[a], T.J. Carrocci[a], and A. A. Hoskins[a]†

Repurposing the "protein-labeling toolkit" for RNA research could be a pragmatic approach for developing new RNA-labeling methods. We have evolved an RNA aptamer that tightly binds benzylguanine (bG), the key ligand for the protein SNAP-tag. The aptamer tightly binds bG fluorophores and can be purified from cellular RNA with bG agarose under native conditions.

One of the defining features of a chemical approach for studying biology is the specific targeting of small molecules to biological targets (tagging). Protein biochemists have an excellent toolkit for tagging proteins of interest for studying cellular localization, biophysical characterization, or for use in protein purification. Tagging proteins with green fluorescent protein (GFP) or derivatives has been used for a wide-range of microscopy experiments[1] in addition to detection or isolation of GFP-tagged proteins with anti-GFP antibodies.[2] More recently, a number of protein tags have been developed that can be used to target small molecules to proteins of interest. Some of these tags, like the SNAP and Halo tags, are self-labelling, meaning that they covalently react with a suicide substrate such as benzylguanine (bG) or chloroalkane derivatives for the SNAP or Halo tag, respectively.[3,4] Other protein tags can tightly bind small molecules such as the strong interaction between dihydrofolatereductase and trimethoprim conjugates (Ligand Link).[5] Finally, some protein tags act as substrates for other enzymes which can be used to install small molecules (i.e., sortase, acyl carrier protein, or biotin ligase).[6] All of these tags have found use both *in vitro* and *in vivo* and a significant amount of work has provided information on the cell permeability and non-specific interactions of the small molecule substrates of these protein tags.[7] Many of the small molecules that show desirable properties (*e.g.*, cell permeability and low non-specific binding) are commercially available and widely used by protein biochemists.

In contrast with protein labelling technologies, RNA labelling is an emerging field. By far the most common methods for labelling RNAs in cells are hybridization of fluorescent oligos (fluorescence *in situ* hybridization)[8] or targeting GFP-tagged proteins to the RNA with specific RNA sequences (e.g., MS2-labeling)[9,10]. More recently RNA aptamers have been developed that tightly bind small molecules, including fluorophores and antibiotics (e.g., the Spinach, Mango, or tobramycin aptamers)[11-13]. Some of these aptamers have been used to image RNAs inside cells or to purify RNA complexes. However, many of the small molecule derivatives that interact with these RNA aptamers are not yet commercially available.

We wondered if commercially available, small molecule tools for protein tagging and already vetted for cell permeability and non-specific interactions could be repurposed as RNA labelling reagents. Specifically we sought to develop RNA aptamers that tightly associate with small molecule protein tags. Here we have carried out Systematic Evolution of Ligands by Exponential Enrichment (SELEX)[14] to evolve a tight-binding RNA aptamer against bG—the key functional group recognized by the protein-based SNAP tag (**Figure 1A**). Functional and structural probing of the evolved aptamer revealed that it binds with high affinity to bG. guanine, related metabolites are poor competitors for bG binding and bG appears to induce a change in RNA aptamer structure. Finally, the aptamer can bind several commercially available fluorescent derivatives of bG and be purified from a complex pool of total cellular RNA using commercial reagents—key features for development and wide-spread adoption of an RNA affinity tag.

To prepare RNAs for SELEX, we created a double stranded DNA transcription template library by PCR amplification of an oligonucleotide containing two, 24-nucleotide (nt) regions of random sequence (**Figure S1**). The random sequence regions were separated by a small, 12 nt constant region that would encode a hairpin topped by a UUCG tetraloop in the transcribed RNAs. A similar, partially-structured library design has been used to evolve other aptamers and can lead to high binding affinities.[11,15] The final library was predicted to contain approximately 10^{14} unique sequences and was used to prepare RNAs by *in vitro* transcription. Sequencing of samples from this library confirmed that it contained the expected random sequences separated by the constant region.

We then carried out SELEX using two different solid supports in parallel: commercially available bG-derivitized agarose (SNAP-Capture Pull-Down Resin, New England Biolabs) and bG-deritivized magnetic DynaBeads that were prepared by incubation of an amine derivative of bG with *p*-toluene sulfonyl-activated beads. In both cases, RNAs were incubated with the support, eluted with high concentrations of free bG, and reverse transcribed to prepare cDNA for subsequent rounds of selection. We followed the progress

[a] Department of Biochemistry, U. Wisconsin-Madison, 433 Babcock Dr., Madison, WI 53706 USA
† corresponding author, ahoskins@wisc.edu
Electronic Supplementary Information (ESI) available including Materials and Methods, Supplemental Figures, and a Supplementary Table. See DOI: 10.1039/x0xx00000x

the selection by incorporation of [^{32}P] into the RNAs to accurately quantify how much RNA was being retained on the support.

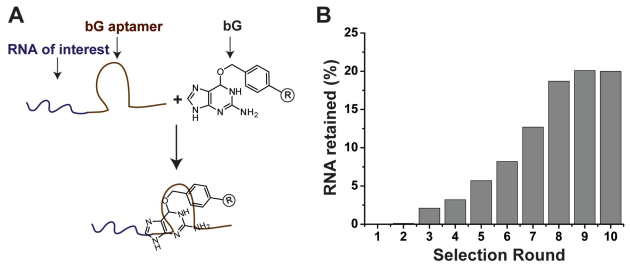


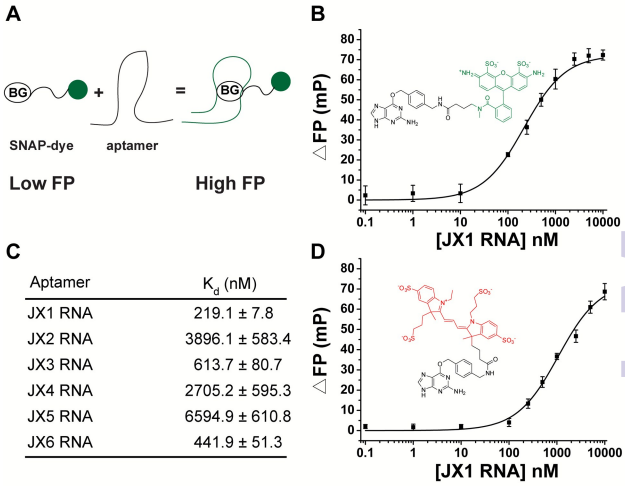
Figure 1. Cartoon of a bG-binding RNA aptamer and results from SELEX. (A) A bG-binding RNA aptamer can be used to target small molecule derivatives of bG to RNAs of interest. R represents a functional group such as a fluorophore or biotin. (B) Quantification of the fraction of RNA retained on bG agarose following each round of SELEX.

Upon quantification of RNA binding during each round of SELEX, we obtained very different results using bG agarose compared with bG Dynabeads. With bG agarose, the fraction of RNA (as measured by [^{32}P] counts per minute) retained on the resin increased from background levels during the first round of SELEX to ~20% after the ninth round (**Figure 1B**). Since no additional enrichment was observed between rounds nine and ten, we did not carry out further rounds of selection. In contrast, we observed no significant enrichment of retained RNAs using bG Dynabeads after ten rounds of selection (data not shown). Nonetheless, we submitted DNAs prepared from reverse transcription of eluted RNAs from both bG agarose and bG Dynabeads for sequencing.

Surprisingly, some of the same RNA sequences were found from samples obtained from both bG agarose and bG Dynabeads (**Table S1**). The JX1 sequence was found in 38% (23 of 60) of sequenced samples obtained from the bG agarose SELEX and 14% of sequenced samples obtained from bG Dynabeads (5 of 36). Since we obtained the same RNA sequence from two different selections, it strongly suggested that the JX1 sequence bound bG. Four other sequences (JX2-5) were obtained from bG agarose that were less abundant than JX1, and of these JX2 was also observed when bG Dynabeads were used. The most common sequence selected using bG Dynabeads (JX6) was present in 25% of those samples (9 of 36) and was not found in the results from the bG agarose selection.

To directly test that the evolved sequences were capable of binding bG independent of the SELEX protocol, we constructed a transcription template encoding the most frequently observed aptamer sequence (JX1) and transcribed the DNA to produce a [^{32}P]-labeled product. We then tested the RNA for specific binding to bG agarose and improved binding to bG agarose when compared with the starting RNA pool. The JX1 aptamer strongly associated with bG agarose, which was able to capture ~70% of the input RNA **(Figure S2A**). The bG agarose resin was ~300-fold more selective for JX1 compared to the mock agarose resin. Additionally, we observed little binding of the initial RNA pool to bG agarose suggesting that only certain RNA sequences are capable of tightly binding bG agarose **(Figure S2B)**. Together these results indicate that JX1 is binding specifically to bG rather than non-specifically associating with the resin.

To measure the affinity of the selected RNA aptamers for bG, we used a fluorescence polarization (FP) assay[16] and the fluorescent bG derivative Atto488-bG (SNAP-Surface 488®, New England Biolabs). In this assay, FP of Atto488-bG would increase when the dye is bound by the aptamer (**Figure 2A**). We transcribed, purified, and measured the binding affinity of aptamers JX1-6 for Atto488-bG. As expected, JX1, the most frequently observed aptamer sequence, also bound Atto488-bG the most tightly (K_d = 219 nM) compared to JX2-6 **(Figure 2B)**. No change in FP was observed when the initial RNA pool was incubated with the fluorophore **(Figure S3).** The other aptamers displayed a range of binding affinities from 0.4-6.6 µM (**Figure 2C and S4)**. Based on these results, we selected JX1 for further analysis.



Aptamer	K_d (nM)
JX1 RNA	219.1 ± 7.8
JX2 RNA	3896.1 ± 583.4
JX3 RNA	613.7 ± 80.7
JX4 RNA	2705.2 ± 595.3
JX5 RNA	6594.9 ± 610.8
JX6 RNA	441.9 ± 51.3

Figure 2. FP assay for measuring affinity of evolved aptamers to fluorescent bG derivatives. (A) Cartoon depicting the FP assay. An increase in FP of the bG derivative is expected upon aptamer binding. (B) Results from FP analysis of Atto488-bG (shown) binding to JX1. (C) Results of FP analysis of Atto488-bG binding to aptamers JX1-6. (D) Results from FP analysis of Dy549-bG (shown) binding to JX1 showing a calculated K_d of 1070±160 nM. In (B) and (C), each data point represents the average of three measurements ± S.D.

To test how well JX1 binds bG derivatives other than Atto488-bG we carried out the FP assay with a bG derivative containing a structurally and spectrally different fluorophore (Dy549-bG, SNAP-Surface 549®, New England Biolabs). We were able to observe a change in FP upon incubation of Dy549-bG with the JX1 aptamer; however, analysis of the binding data revealed a decrease in binding affinity (K_d= 1070 nM, **Figure 2D**). The decrease in affinity may be due to unfavorable interactions between the tetrasulfonated cyanine fluorophore of Dy549 and the RNA. Less negatively charged molecules, like the disulfonated fluorone derivative Atto488-bG, may interact more favorably with RNA and lead to higher affinities.

To test how selective JX1 is for bG over other guanine metabolites we used a FP competition assay[17]. In this experiment, we would predict that competitors would result in an increase in free Atto488-bG and a corresponding loss in FP (**Figure 3A)**. In all cases, we introduced the competitor and Atto488-bG simultaneously to JX1. As a proof-of-concept, non-fluorescent bG (20µM) was able to effectively compete with Atto488-bG (50 nM) and reduce the observed FP to near background levels (**Figure 3B**). This result also indicates that bG lacking PEG linkers or functional groups is able to be bound by JX1, in agreement with isolation of RNAs during SELEX by bG elution.

We next analyzed changes in FP upon addition of guanine, guanosine, guanosine 5' –monophosphate (GMP), or guanosine 5' triphosphate (GTP) **(Figure 3B)**. Unlike with bG, we observed no

change in fluorescence polarization with any of these metabolites when they were introduced in 400-fold excess of the Atto488-bG. At even higher concentrations of competitor (100 µM), we observed a slight decrease in FP with guanine but not the other molecules. GMP and GTP did not change the FPeven when included at 1 mM concentration (20,000-fold excess). At 10 mM GMP and GTP we did observe a loss in FP; however, at these high concentrations it is possible that the changes observed could also be due to quenching of the fluorophore by GMP or GTP[18]in addition to competition for the ligand binding site. From these data, we conclude that common guanine metabolites are poor competitors of Atto488-bG for binding to JX1. These data also may indicate that JX1 recognizes bG through interaction with the benzyl moiety and/or through interactions with the guanine N9 position, which is available for hydrogen bonding in bG but part of the glycosidic linkage in guanosine, GMP, and GTP.

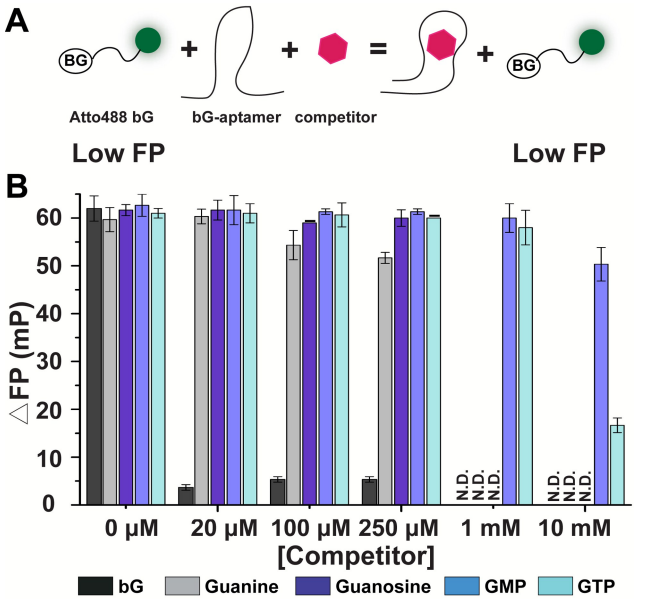


Figure 3. FP competition assay for monitoring specificity of JX1 for bG. (A) Cartoon showing how an effective competitor of Atto488-bG for JX1 binding prevents binding of the RNA to the fluorophore results in low FP. (B) Results from the competition assay for selected guanine-containing molecules. High ΔFP values result from Atto488-bG binding to JX1 and lack of inhibition from the competitor.Each bar graph represents the average from three separate experiments and error bars represent ± S.D. N.D., Not Determined.

Structural analysis JX1 by Mfold[19]suggests that the RNA can fold into a complex structure containing multiple stem loops (**Figure 4A**). To further analyze the structure, we carried out enzymatic cleavage assays using RNase T1[20] both in the presence and absence of bG (**Figure 4B**). Analysis of the T1 digest showed significant changes between the free aptamer and after complex formation with bG. The data support a conformational change occurring in the RNA that leads to shielding of guanines in putative stem loops I and III after bG binding. In contrast, guanines located between these regions become more susceptible to RNase T1 cleavage. The latter observation was corroborated by increased cleavage at U64 during in line probing (data not shown).

Interestingly, we also observed changes in JX1 in the presence of bG that correspond to locations found within the primer binding site used in the SELEX protocol. This suggests that these regions may also be important for the structure of JX1 and bG binding. In support of this, we truncated JX1 to remove 30 nt from the 5' and 13 nt from the 3' end of the RNA (JX1, Δ1-30, Δ87-100) and analyzed Atto488-bG binding by FP. The results showed little evidence for Atto488-bG association **(Figure 4C)**—confirming our prediction that regions at the very 5' and 3' ends of JX1 are important for aptamer function.

We next attempted to minimize JX1 by removal of the constant region encoding a UUCG-tetraloop that was present in the initial DNA library (**Figure S1**). Deletion of this region resulted in an aptamer (JX1, Δ41-60) that was still functional for Atto488-bG binding. However, the binding affinity was reduced ~4-fold (K_D = 957 nM, **Figure 4C**).This indicates that not only are primer binding regions used in the SELEX protocol essential for binding, but the UUCG tetraloop region also plays an important functional role. In sum, these experiments suggest that JX1 is a conformationally complex and potentially dynamic RNA aptamer.

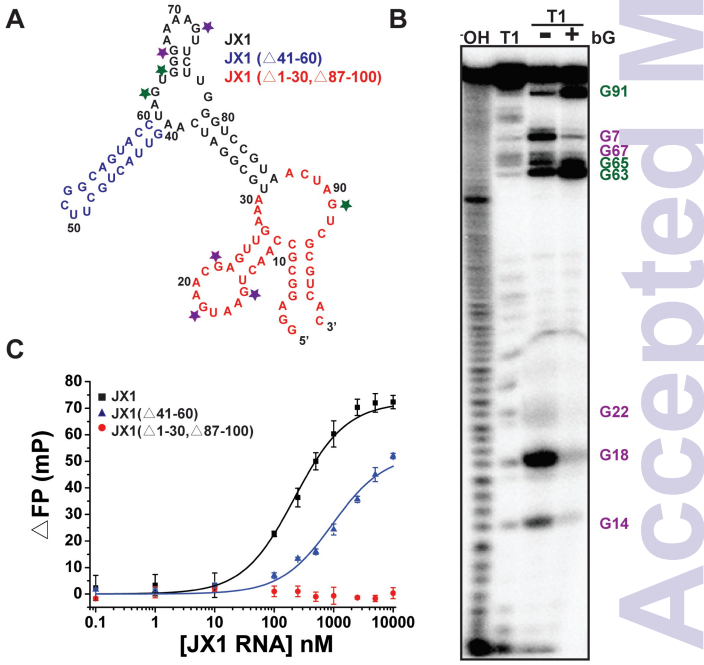


Figure 4. Structural characterization of JX1. (A) Secondary structure prediction of JX1 based on Mfold analysis. Regions truncated in efforts to minimize JX1 are shown in blue and red. Purple and green stars represent nucleotides that showed decreased or increased RNase T1 cleavage after addition of bG. (B) Results from RNase T1 digestion of JX1. Lane 1, RNA ladder produced by alkaline hydrolysis. Lane 2, RNase T1 digestion under denaturing conditions. Lanes 3 and 4, RNase T1 digestion under native conditions in the presence of absence of bG. Positions of increased or decreased RNase T1 cleavage seen upon addition of bG are noted. (C) FP assay of minimized JX1 RNAs (red and blue) compared with full-length JX1 (black). No Atto488-bG binding was observed with the JX1(Δ1-30, Δ87-100) RNA (red). The JX1(Δ41-60) RNA was able to bind Atto488-bG (K_d = 957±52).

Two key features of the JX1 aptamer are the commercial availability of bG-deritivitized reagents and the ability to elute the JX1 aptamer from bG agarose under native conditions without RNA unfolding. We wondered if these features could be used to purify JX1 from a complex RNA mixture. To test this hypothesis, we combined JX1 (10 nM) with unfractionated yeast cellular RNA (total yeast RNA, 1 µg). The resulting RNA mixture was then heated to

denature the RNAs and allowed to cool before incubation with commercially available bG agarose. The resin was then placed into a column, and the flowthrough was collected. Bound RNAs were eluted by first washing the column with five volumes of buffer and then eluting with a solution of commercially available bG (5 mM). RNAs present at each stage were then visualized by denaturing polyacrylamide gel electrophoresis and ethidium bromide staining.

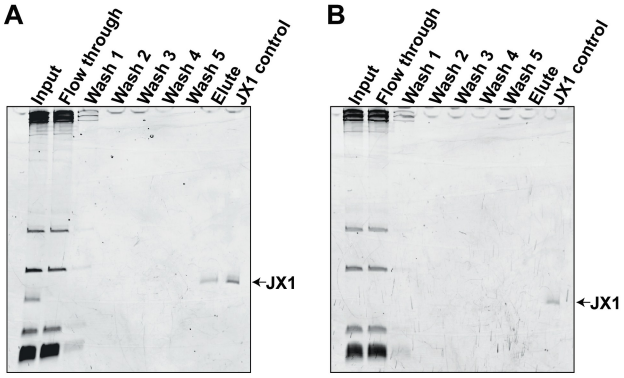


Figure 5. Purification of JX1 from total yeast RNA. (A) JX1 can be purified and eluted with bG under native conditions in the presence of cellular RNAs. (B) In a control experiment, no band of similar size to JX1 could be observed in either the input or eluate if JX1 was omitted from the reaction mixture.

Under these conditions, JX1 was readily purified from total yeast RNA and eluted from bG agarose without the use of a denaturant (*e.g.,* urea, formamide, or EDTA). Recovery of JX1 was ~50% based upon analysis of band intensities from the input and eluate samples (**Figure 5A**). No band of similar size to JX1 was present in either the input or eluate from a control experiment in which the aptamer was not included with the yeast total RNA (**Figure 5B**). These results reveal not only can JX1 be purified from cellular RNAs under native conditions but also that JX1 can fold into a structure capable of binding bG agarose after first being denatured in the presence of cellular RNAs.

In conclusion we have evolved an RNA aptamer that specifically and tightly binds bG ligands that currently in wide-spread use as substrates for the protein SNAP tag. We have shown that these ligands can be used to both target fluorphores to the aptamer as well as purify the aptamer from a complex mixture of nucleic acids under native conditions. Development of a bi-functional small molecule library—whose members can be used to label either proteins or RNAs—could represent a valuable addition to the chemical biology toolkit. Further engineering of the JX1 aptamer may lead to other, more diminutive bG binding or reactive aptamers that can function analogously to the protein SNAP tag.

We thank the Dan Stevens and the Biophysical Instrumentation Facility (BIF, Dept of Biochemistry, U. Wisconsin-Madison) for instrument access.JX, TJC, and AAH are supported by a Beckman Young Investigator award to AAH and startup funds from the Dept. of Biochemistry, WARF, and the U. Wisconsin-Madison. AAH is also supported by a Shaw Scientist Award from the Greater Milwakee Foundation. TJC is also supported by the NIH Biotechnology Training Program at U. Wisconsin-Madison 5T32GM08349) and a William H. Peterson Fellowship.

References

1. D.M. Chudakov, M.V. Matz,S. Lukyanovand K.A. Lukyanov, *Physiol. Rev.*,2010,**90,**1103.
2. R.Zhuang, Y.Zhang, R. Zhang, C.J.Song, C.H.Song, K. Yang, A. Yang and B.Q. Jin, *ProtExpr&Pur.*, 2008, **59**, 138.
3. A. Juillerat, T. Gronemeyeret, A. Keppler, S. Gendreizig, H. Pick. H. Vogel and K. Johnsson, *Chem Biol.,* 2003,**10**, 313.
4. G.V.Los and K. Wood,*MethodsMol Biol.,* 2007, **356**, 195.
5. L.W. Miller, Y.F. Cai, M.P.Sheetz and V.W. Cornish. *Natur Method.* 2005,**2**, 255.
6. M. Rashidian, J.K. Dozier and M.D. Distefano, *Bioconjugate Chemistry*, 2013,**24**, 1277.
7. P. J. Bosch, I.R. CorrêaJr, M. H. Sonntag, J. Ibach, L. Brunsveld, J. S. Kanger and V. Subramaniam. *Biophys. J.*, 2014, **107**, 803.
8. A.M. Femino, F.S. Fay, K. Fogarty and R.H. Singer, *Science*. 1998,**280,** 585.
9. E. Bertrand, P. Chartrand, M. Schaefer, S.M. Shenoy, R.H Singer and R.M. Long. *Mol. Cell.,*1998.**2**.437.
10. T. J. Carrocci and A. A. Hoskins, *Analyst*, 2014, **139**, 44.
11. J.S. Paige, K.Y. Wu, S.R. Jaffrey, *Science*, 2011,**333**, 642.
12. E.V. Dolgosheina, S.C. Jeng, S.S. Panchapakesan, R.Cojoc, P.S. Chen, P.D. Wilson, N. Hawkins, P.A. Wiggins and P.J. Unrau. *ACS Chem. Biol.,* 2014, **9,** 2412.
13. I. Shin, J. Ray, V. Gupta, M. Ilgu, J. Beasley, L. Bendickson, S. Mehanovic,G.A. Kraus and H.M. Nilsen, *Nucleic Acids Res.,* 2014,**42**,e90.
14. C. Tuerk and L. Gold, *Science*, 1990, 249, 505.
15. J.H.Davis, and J.W.Szostak, *Proc. Natl.Acad.Sci. U.S.A.,* 2002 **99,** 11616.
16. J.M. Pagano, C.C. Clingman and S.P. Ryder, *RNA,* 2010, **17,** 14
17. X. Geng, D.P. Zhang, H.L. Wang and Q. Zhao. *Anal Bioanal Chem*., 2013, **405**, 2443.
18. I.Nazarenko, R. Pires, B. Lowe, M. Obaidy and A. Rashtchian, *Nucleic Acid Res*., 2002, **30**, 2089.
19. M. Zuker. *Nucleic Acid Res*., 2003, **31**, 3406.
20. M.F. Carey, C.L. Peterson, S.T. Smale, *Cold Spring HarbProtoc*. 2013, **3,**doi: 10.1101/pdb.prot071910.