Genetically encoded RNA nanodevices for cellular imaging and regulation

Qikun Yu,† Kewei Ren† and Mingxu You D

Nucleic acid-based nanodevices have been widely used in the fields of biosensing and nanomedicine. Traditionally, the majority of these nanodevices were first constructed in vitro using synthetic DNA or RNA oligonucleotides and then delivered into cells. Nowadays, the emergence of genetically encoded RNA nanodevices has provided a promising alternative approach for intracellular analysis and regulation. These genetically encoded RNA-based nanodevices can be directly transcribed and continuously produced inside living cells. A variety of highly precise and programmable nanodevices have been constructed in this way during the last decade. In this review, we will summarize the recent advances in the design and function of these artificial genetically encoded RNA nanodevices. In particular, we will focus on their applications in regulating cellular gene expression, imaging, logic operation, structural biology, and optogenetics. We believe these versatile RNA-based nanodevices will be broadly used in the near future to probe and program cells and other biological systems.

1. Introduction

Naturally existing DNAs and RNAs are polymer chains of nucleotides, which are composed of only four bases: thymine/uracil, guanine, adenine, and cytosine. While, due to the diverse order of nucleotides, highly specific Watson–Crick base pairing, flexible design, and self-assembly property, nucleic acids have been widely used as a promising building material for various nanostructures and devices.1–4 In recent years, a diversity of two-dimensional or three-dimensional nucleic acid-based nanodevices have been constructed with unique features of structural programmability, spatial addressability, controllable length, size and shape, and easy synthesis and functionalization.5,6 These nanodevices, especially DNA-based ones, have been extensively applied in structural biology, bionanotechnology, in vitro diagnostics, cell membrane analysis, and nanomedicine.5–11

Even though powerful, several concerns have been raised for the intracellular and in vivo applications of these nanodevices. For example, the majority of nucleic acid-based nanostructures and devices are in vitro prepared using chemically or enzymatically synthesized DNA/RNA oligonucleotides. For...
intracellular applications, these nanodevices have to be first delivered into cells. Even though several DNA nanodevices have been successfully developed for intracellular imaging, unfortunately, more general, highly efficient and non-invasive cellular delivery of nucleic acids is not always feasible.\textsuperscript{12-15} Meanwhile, the enzymatic degradation of DNA/RNA, highly complex cellular environment, and potential cytotoxicity of synthetic compounds have made it even more challenging to apply artificial nucleic acid-based nanodevices for cellular studies.\textsuperscript{16,17} In addition, the relatively low stability of large DNA nanostructures, e.g., DNA origami, at physiological Mg\textsuperscript{2+} concentrations also restricts their applications inside cells.\textsuperscript{16}

Compared with synthetic DNA nanodevices, functional RNA molecules can be genetically encoded and directly synthesized inside living cells using natural transcription machinery. A variety of RNA nanodevices exist in nature, including riboswitches, ribozymes, transfer RNAs, ribosomal RNAs, etc. These natural RNA devices play important roles in cellular functions by regulating metabolite recognition, RNA processing, and gene expression.\textsuperscript{18-20} Inspired by these natural RNA nanodevices, RNA nanotechnology has recently emerged to construct artificial functional devices for cellular imaging and regulation.\textsuperscript{21-23} These RNA-based artificial devices can also be genetically encoded and continuously produced inside cells. The cellular expression of these nanodevices can be maintained at a constant level for a long period of time and across generations. The function of these RNA nanodevices can also be activated in either specific target cells or throughout the whole cell populations.\textsuperscript{24,25} As a result, these genetically encoded RNA devices provide an elegant solution to the problems with traditional synthetic nucleic acid-based tools in cellular deliveries and maintaining cellular concentrations and functions.

In this review, we will discuss the recent progress in the design, construction, and application of artificial genetically encoded RNA nanodevices. These artificial nanodevices can be rationally and programmably designed, providing a modular platform for intracellular applications. On one hand, combining target-recognition functions with on-demand regulatory activities, artificial RNA nanodevices have been used for cellular gene regulation.\textsuperscript{21,22} On the other hand, genetically encoded RNA devices can also be applied as sensors for imaging and detecting various target analytes in living cells.\textsuperscript{23} More recently, even further advanced RNA nanodevices have emerged to regulate cellular protein and RNA networks, construct cellular logic circuits, and optogenetically control RNA functions. By summarizing the design principles and features of existing RNA-based nanodevices, we hope this review will potentially inspire new structures and functions of these exciting molecular machines in living systems.

2. Genetically encoded RNA nanodevices for gene regulation

In nature, both coding and non-coding RNAs play key roles in controlling cellular gene expression.\textsuperscript{26,27} Naturally existing regulatory RNAs are ubiquitous, but also highly conserved and sophisticated, which makes them hard to be modulated or altered for designable functions. Nowadays, artificially designed RNA nanodevices can also fold and react with protein partners to achieve gene regulation, but in contrast, the structure and function of these artificial RNA devices can be better predicted and controlled based on computational models, and software such as Mfold and NUPACK are freely available online.\textsuperscript{28,29} The development of artificial RNA nanodevices to sense and regulate genes has become an important research area in synthetic biology. Compared with protein-based transcriptional and translational regulation,\textsuperscript{30} the use of RNAs has several advantages, including their easily predictable base-pairing interactions, dynamic binding-induced conformational changes, and the ability of systematically evolving new ligand-recognition units, e.g., aptamers.\textsuperscript{31} To date, a couple of RNA-based designs have been developed into powerful gene regulation devices both \textit{in vitro} and \textit{in vivo}.\textsuperscript{32-36} In this section, we will focus on artificially designed genetically encoded RNA nanodevices that have been validated inside living cells for gene expression and regulation.

2.1 Toehold reaction-based RNA nanodevices

One natural mechanism in achieving post-transcriptional gene regulation is by blocking the ribosomal access to the initiation site using either antisense small regulatory RNAs or mRNA-binding proteins.\textsuperscript{37} Mimicking this natural regulation mechanism, the Collins group reported an artificial riboregulator device in 2004 that can either repress or activate translation in \textit{Escherichia coli} (\textit{E. coli}).\textsuperscript{38} This engineered riboregulator consisted of two parts: a \textit{cis}-repressed mRNA (crRNA) and a \textit{trans}-activating RNA (taRNA) (Fig. 1a). The designed self-folding in the 5’-untranslated region of the crRNA kept the ribosome binding site (RBS) in a duplex formation and blocked it from recognizing the 30S ribosomal subunit. As a result, protein translation was inhibited. On the other hand, the taRNA was
designed to hybridize with the stem and loop region in the crRNA. The resulting RNA duplex formation unfolded the crRNA, exposed the RBS region and permitted translation. A maximum of 19-fold increase in the GFP expression was observed in *E. coli* cells after activating the taRNA.38

Following this initial study, a further optimized riboregulator with larger fold of gene activation and better ability of expressing multiple genes was demonstrated in *E. coli*.39 Meanwhile, with minimal leakage, RNA riboregulators have been used to develop a programmable kill switch for bacteria.40 Based on a similar design principle, several other RNA riboregulators have been developed as well, which are again mostly tested within *E. coli* cells.41,42 Even though these RNA riboregulators are able to regulate gene expression inside cells, these systems still suffer from a modest dynamic range, low specificity and orthogonality, and limited choice of sequences. The moderate dynamic range mainly stems from thermodynamically and kinetically unfavorable loop region-mediated interactions,43,44 while limitations in sequence and low orthogonality come from the requirement of a double-stranded RBS region formation in the crRNA.

To overcome these challenges, an elegant RNA nanodevice was developed by the Yin group in 2014, which was named toehold switches.45 These toehold switches consisted of a switch RNA and a trigger RNA (Fig. 1b). Unlike traditional RNA riboregulators, the RBS and start codon (AUG) regions in the switch RNA were completely unpaired, and as a result, the sequence of the trigger RNA is no longer constrained. Meanwhile, by replacing the “loop–loop” or “loop–linear” interaction with a toehold-mediated “linear–linear” interaction between unstructured RNAs, the accessibility of the trigger RNAs is also increased. With thermodynamically and kinetically more favorable toehold-mediated strand displacement reactions, these toehold switches can provide a much larger dynamic range (on average >400-fold) and better orthogonality in *E. coli* cells. These properties of toehold switches have already been compared with those of optimized protein-based gene regulators.45 These toehold switches have been further integrated into bacterial genomes to regulate endogenous genes and incorporated into cell-free paper-based platforms for *in vitro* diagnostics.46 More recently, toehold switches have also been engineered into various mammalian cell lines, including HEK293, HeLa and MDA-MB-231, for detecting microRNAs.47

Toehold reaction-based RNA nanodevices can not only regulate genes at the translational level, but also at the transcriptional level.48,49 For example, the Lucks group created a small transcription activating RNA (STAR) nanodevice in 2015 to regulate bacterial transcription in *E. coli*.48 This STAR system was composed of a transcription terminator-containing gene and a STAR antisense small RNA (Fig. 1c). Inspired by the naturally existing pT181 transcriptional attenuator, an intrinsic terminator hairpin was designed to fold within the upstream of the regulated gene. The formation of this terminator caused RNA polymerase to terminate transcription before reaching the gene of interest. In the presence of the STAR antisense RNA, the terminator hairpin could be opened through the toehold-mediated strand displacement reaction, which allowed the transcription elongation of the target gene. Based on this design principle, orthogonal pairs of STARS were developed with up to 94-fold gene activation.48 By further incorporating a computational design approach, highly efficient and orthog-
onal STARs have been engineered with maximally ~9000-fold gene activation in *E. coli*. Similarly, toehold reaction-based RNA nanodevices can be used to transcriptionally inhibit bacterial gene expression as well.

### 2.2 Riboswitch- and ribozyme-based RNA nanodevices

For decades, proteins have been considered as the only cellular component that could specifically recognize and respond to small molecules. This observation changed after the discovery of two kinds of natural RNA devices that can perform genetic regulation (i.e., riboswitches) or catalytic reactions (i.e., ribozymes). These naturally evolved RNA elements have been further engineered for specific gene regulation inside cells.

Since the first report of natural RNA riboswitches that can bind vitamin derivatives to regulate cellular vitamin synthesis in 2002, plenty of riboswitches have been discovered in bacteria, archaea, plants, and fungi to recognize various target molecules including protein cofactors, nucleotides, amino acids, sugars, and ions. These RNA riboswitches are composed of two components: a target-sensing aptamer domain and an expression platform (Fig. 2a). The target binding to the aptamer domain induces a conformational change in the expression platform, leading to the regulation of downstream genes. RNA riboswitch-based gene regulation can be achieved at both translational and transcriptional levels. In the case of translation control, target binding-induced RNA structural changes is often coupled with the sequestering or release of the RBS region, similar to that of toehold switches.

To regulate transcription, similar to the STAR design, natural transcriptional attenuators are normally required in the function of these RNA riboswitches.

With the help of RNA riboswitches, a number of cellular targets, especially metabolites, can now be incorporated for gene regulation. However, for a majority of synthetic compounds and many cellular components, natural RNA riboswitch partners have not yet been identified. These compounds, especially the synthetic ones, can be potentially highly useful in generating bioorthogonal gene regulation units in synthetic biology.

To identify specific RNA sequences for “any” given target molecule, Systematic Evolution of Ligands by Exponential enrichment (SELEX) is often used for the *in vitro* selection of RNA aptamers. Like natural RNA riboswitches, some of these *in vitro*-identified RNA aptamers can also undergo conformational changes after binding with the target. As a result, these synthetic aptamers can be similarly placed in the 5′ untranslated region of an mRNA to regulate genes. Interestingly, the initial successful attempt of these synthetic aptamer-based nanodevices was actually achieved in 1998 even before the discovery of natural RNA riboswitches.

In this case, aptamers that target Hoechst dyes were used to downregulate gene expression in Chinese hamster ovary cells. Similarly, the ability of *in vitro*-selected tetracycline and neomycin aptamers in gene regulation was also proved in yeast. Besides these, synthetic theophylline-targeting RNA riboswitch is one of the most widely used RNA nanodevices for intracellular gene regulation. This is because theophylline exhibits great cell permeability and bioorthogonality; in addition, the corresponding aptamer can recognize theophylline with good specificity and binding affinity. After binding with theophylline, a predictable RNA conformational change is induced as a result.

Even though much success has been achieved using synthetic RNA aptamers generated through *in vitro* SELEX, concerns over the *in vivo* selectivity and folding patterns of these nanodevices have influenced the broad usage of this new gene regulation platform. To overcome this challenge, several attempts have been made based on directed mutagenesis to reengineer natural RNA riboswitches for the binding of non-natural target ligands.

These reengineered synthetic riboswitches were proved to function orthogonally to their original targets, which provided a promising alternative approach to regulate intracellular genes.

With the increasing need for reliable and predictable nanodevices for gene regulation, another type of natural functional RNA molecule, ribozymes, has also been popularly used in engineering synthetic RNA devices. Natural RNA ribozymes are mostly involved in the catalytic processing of intron excision. By fusing riboswitches or *in vitro* selected RNA aptamers with these ribozyme molecules, the resulting RNA nanodevices, termed ribozyme switches or aptazymes, are able to regulate gene expression in the presence of cognate target ligands (Fig. 2b). The most commonly used ribozyme for this purpose is the hammerhead ribozyme, which can perform *cis*- or *trans*-RNA cleavage once a three-way-junction catalytic core structure is formed. Aptamers are normally fused into one stem junction of the hammerhead ribozyme. In the absence of the

---

**Fig. 2** (a) Schematic of a riboswitch-based RNA nanodevice for the gene regulation during prokaryotic protein synthesis. The binding of a target ligand induces the RNA conformational change, exposes the ribosome binding site (RBS) for the protein synthesis. (b) Schematic of a hammerhead ribozyme-based RNA nanodevice for the gene regulation. The binding of a target ligand induces the folding and catalytic function of the hammerhead ribozyme. As a result, the RBS region is released and exposed to start the protein synthesis.
target, this stem region is unfolded, resulting in minimal cleavage. The target binding to the aptamer region induces the formation of the three-way-junction catalytic core and activates the ribozyme cleavage around the RBS region. It will then lead to the regulated protein translation. Using this design principle, hammerhead ribozyme-based RNA nanodevices have been activated inside cells using various target molecules such as thiamine pyrophosphate, theophylline, tetracycline, etc. In addition to the hammerhead ribozyme, other types of ribozymes, e.g., the twister ribozyme, could also be similarly used for gene regulation. The catalytic function of these ribozymes has been used to improve the sensitivity of RNA nanodevices, while in the meantime, background signal leakage from spontaneous RNA cleavage has to be carefully optimized to achieve a large fold of gene activation/inhibition.

2.3 CRISPR-based RNA nanodevices

Clustered regularly interspaced short palindromic repeats (CRISPR), together with CRISPR-associated protein (Cas), is one of the most powerful systems for genetic manipulation. On the basis of a simple RNA-guided sequence-specific DNA recognition, the CRISPR-Cas system is highly programmable and efficient. With the help of engineered Cas9 protein and small guide RNA (sgRNA), targeted gene editing has been successfully demonstrated in various prokaryotic and eukaryotic cells, tissues, and animals. In addition, without endonuclease activity, the catalytically dead Cas9 (dCas9) protein can also exhibit similar RNA-guided gene targeting property, which is useful for reversible and inducible gene regulation.

Indeed, by designing synthetic RNA nanodevices to regulate sgRNA structures and functions, highly efficient and modular control of the CRISPR-Cas9 and CRISPR-dCas9 systems have been achieved. One success in using dCas9/sgRNA for the sequence-specific control of gene regulation, named CRISPR interference (CRISPRi), was achieved by Qi et al. in 2013. Because dCas9 is catalytically inactive, the dCas9/sgRNA complex would continue binding with the cognate DNA region and sterically block the transcription (Fig. 3a). CRISPRi can repress the target gene with up to 1000-fold efficiency in E. coli cells. By adding several different sgRNA sequences simultaneously, CRISPRi can be also used to regulate multiple genes. In addition, by coupling a transcriptional activator with dCas9, CRISPRi can also facilitate target gene activation.

The above-mentioned toehold switches, riboswitch- and ribozyme-based nanodevices can also be introduced into the CRISPRi system to allow gene regulation by different RNA or small molecule triggers. For example, an inducible CRISPR system was developed in 2016 by the Cai and Huang groups. In their design, a tetracycline-binding aptamer was inserted into the 3’ end of the sgRNA. The stem region of the aptamer was designed to hybridize with the guide sequence of the sgRNA to inhibit its binding to the target DNA (Fig. 3b). The addition of tetracycline refolded the sgRNA and led to the desired binding and regulation of the target gene, as demonstrated in the HEK293T cells. By replacing the tetracycline-binding aptamer with other aptamer sequences, such as a theophylline-binding aptamer, different small molecules can also be used to regulate the efficiency of CRISPR-based gene regulation.

These target-binding aptamers can also be inserted in the middle of the sgRNA to achieve more versatile small molecule-controlled CRISPRi. Theophylline- and 3-methylxanthine-binding aptamers have been used to demonstrate the function of these nanodevices for bacterial gene regulation. In another design, the Liu group has engineered a ribozyme-controlled sgRNA by linking a guanine-targeting aptazyme to the 5’ end of the sgRNA (Fig. 3c). The binding of guanine induced the cleavage of the ribozyme and activated the sgRNA for the cognate DNA recognition in the HEK293T cells, while it is worth mentioning that the self-cleavage of ribozyme could lead to some signal leakage in this system.

Besides small molecule-based triggers, nucleic acids can also be used to control the structure of the sgRNA. For example, conditional activation of the sgRNA can be achieved with the addition of a toehold sequence in its 5’ end, sequence of which will also block these DNA-recognition domains in both bacterial and mammalian cells. In the presence of another trigger RNA, the sgRNA sequence can be released for targeted gene regulation. Another rational design strategy of inducible sgRNA was achieved by the Fulga group through the incorporation of a natural RNA-cleaving unit, such as endoribonuclease targeting region or antisense oligonucleotide-mediated RNase H cleavage site. A modular and rapid
control of the CRISPR functions was demonstrated based on this design. Indeed, these synthetic RNA nanodevices have provided an attractive and versatility approach in regulating CRISPR-based genetic modification.

3. Genetically encoded RNA nanodevices for intracellular imaging

3.1 Fluorogenic RNA aptamer-based imaging tags

In addition to regulating gene expression, another promising application of genetically encoded RNA nanodevices is cellular imaging and detection of various target analytes. Traditionally, fluorescent protein (FP)-based reporters have been commonly used to construct protein- or RNA-based sensors for intracellular imaging.\textsuperscript{95} For example, based on the specific binding between an MS2 RNA hairpin and MS2 bacteriophage coat protein, RNA targets that are tagged with multiple copies of MS2 RNA can be imaged in living cells with FP-fused MS2-binding proteins.\textsuperscript{96} FP-based RNA nanodevices have also been developed to real-time monitor target analytes in living systems.\textsuperscript{25,97} However, there are still several challenges in applying these FP reporter-based RNA nanodevices. First of all, the large molecular weight of FP may interfere with the location and cellular functions of the nanodevice and target analytes.\textsuperscript{98,99} In addition, the limited choice of orthogonal RNA–protein binding pairs make it difficult for multiplex or programmable detection. Moreover, the high background fluorescence and limited dynamic range of existing FP-based RNA nanodevices have further prevented their wide applications.\textsuperscript{100}

Compared with FP-based reporters, it is more convenient and desirable to have genetically encoded RNA-based fluorescent reporters for these RNA nanodevices. Indeed as a result, fluorogenic RNA aptamers have been prepared. These RNA aptamers can selectively bind to small molecule chromophores and activate the corresponding fluorescence signal.\textsuperscript{101} A number of fluorogenic RNA aptamer/chromophore pairs with different spectral and biophysical properties have been developed recently, including so-called Spinach,\textsuperscript{102} Broccoli,\textsuperscript{103} Mango,\textsuperscript{104} Corn,\textsuperscript{105} Pepper,\textsuperscript{106} DNB,\textsuperscript{107} SRB-2,\textsuperscript{108} etc. (Table 1 and Fig. 4). With a high signal-to-background ratio, easy programmability, and small size, these fluorogenic RNA aptamers have started to being used as a versatile reporting unit in genetically encoded RNA nanodevices for intracellular imaging.\textsuperscript{31}

### Table 1  Spectral and biophysical characteristics of commonly used fluorogenic RNA aptamers

<table>
<thead>
<tr>
<th>RNA aptamer</th>
<th>Fluorophore</th>
<th>$K_D$ (nM)</th>
<th>$E_x/E_m$ (nm)</th>
<th>$\epsilon$ (M$^{-1}$ cm$^{-1}$)</th>
<th>$\Phi$</th>
<th>Length (nt)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinach</td>
<td>DFHBI</td>
<td>540</td>
<td>469/501</td>
<td>24 300</td>
<td>0.72</td>
<td>98</td>
<td>102</td>
</tr>
<tr>
<td>Spinach2</td>
<td>DFHBI-1T</td>
<td>560</td>
<td>482/505</td>
<td>31 000</td>
<td>0.94</td>
<td>95</td>
<td>174</td>
</tr>
<tr>
<td>Spinach2</td>
<td>DFHBI</td>
<td>530</td>
<td>447/501</td>
<td>22 000</td>
<td>0.72</td>
<td>95</td>
<td>174</td>
</tr>
<tr>
<td>Spinach2</td>
<td>DFHBI-2T</td>
<td>1300</td>
<td>500/523</td>
<td>29 000</td>
<td>0.12</td>
<td>95</td>
<td>174</td>
</tr>
<tr>
<td>Spinach2</td>
<td>DFHBI-CM</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>95</td>
</tr>
<tr>
<td>Broccoli</td>
<td>DFHBI-1T</td>
<td>360</td>
<td>472/507</td>
<td>29 600</td>
<td>0.94</td>
<td>49</td>
<td>103</td>
</tr>
<tr>
<td>Broccoli</td>
<td>BI</td>
<td>51</td>
<td>470/505</td>
<td>33 600</td>
<td>0.67</td>
<td>49</td>
<td>113</td>
</tr>
<tr>
<td>Red Broccoli</td>
<td>DFHO</td>
<td>206</td>
<td>518/582</td>
<td>35 000</td>
<td>0.34</td>
<td>49</td>
<td>105</td>
</tr>
<tr>
<td>Orange Broccoli</td>
<td>DFHO</td>
<td>230</td>
<td>513/562</td>
<td>34 000</td>
<td>0.28</td>
<td>49</td>
<td>105</td>
</tr>
<tr>
<td>Red Broccoli</td>
<td>OBI</td>
<td>23</td>
<td>541/590</td>
<td>47 300</td>
<td>0.67</td>
<td>54</td>
<td>126</td>
</tr>
<tr>
<td>Corn</td>
<td>DFHO</td>
<td>70</td>
<td>505/545</td>
<td>29 000</td>
<td>0.25</td>
<td>36</td>
<td>105</td>
</tr>
<tr>
<td>Mango</td>
<td>TO1-biotin</td>
<td>3.6</td>
<td>510/535</td>
<td>77 500</td>
<td>0.16</td>
<td>29</td>
<td>104</td>
</tr>
<tr>
<td>Mango II</td>
<td>TO1-biotin</td>
<td>0.7</td>
<td>510/535</td>
<td>77 000</td>
<td>0.2</td>
<td>29</td>
<td>176</td>
</tr>
<tr>
<td>Mango III</td>
<td>TO1-biotin</td>
<td>5.6</td>
<td>510/535</td>
<td>77 000</td>
<td>0.56</td>
<td>29</td>
<td>176</td>
</tr>
<tr>
<td>Mango IV</td>
<td>TO1-biotin</td>
<td>11.1</td>
<td>510/535</td>
<td>77 000</td>
<td>0.42</td>
<td>29</td>
<td>176</td>
</tr>
<tr>
<td>DIR2s-Apt</td>
<td>DIR-pro</td>
<td>252</td>
<td>600/658</td>
<td>164 000</td>
<td>0.33</td>
<td>57</td>
<td>177</td>
</tr>
<tr>
<td>DIR2s-Apt</td>
<td>OTB-SO3</td>
<td>662</td>
<td>380/421</td>
<td>73 000</td>
<td>0.51</td>
<td>57</td>
<td>177</td>
</tr>
<tr>
<td>Chili</td>
<td>DMHBI      *</td>
<td>63</td>
<td>413/542</td>
<td>21 000</td>
<td>0.4</td>
<td>52</td>
<td>178</td>
</tr>
<tr>
<td>Chili</td>
<td>DMHBO      *</td>
<td>12</td>
<td>456/592</td>
<td>22 000</td>
<td>0.1</td>
<td>52</td>
<td>178</td>
</tr>
<tr>
<td>SiRA</td>
<td>SiR</td>
<td>430</td>
<td>649/662</td>
<td>86 000</td>
<td>0.98</td>
<td>46</td>
<td>179</td>
</tr>
<tr>
<td>Pepper485</td>
<td>HBC485</td>
<td>8</td>
<td>443/485</td>
<td>49 100</td>
<td>0.42</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper497</td>
<td>HBC497</td>
<td>6.7</td>
<td>43/597</td>
<td>54 700</td>
<td>0.57</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper508</td>
<td>HBC508</td>
<td>27</td>
<td>458/508</td>
<td>42 500</td>
<td>0.3</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper514</td>
<td>HBC514</td>
<td>12</td>
<td>458/514</td>
<td>44 100</td>
<td>0.45</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper525</td>
<td>HBC525</td>
<td>3.8</td>
<td>491/525</td>
<td>74 100</td>
<td>0.7</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper530</td>
<td>HBC530</td>
<td>3.5</td>
<td>485/530</td>
<td>65 300</td>
<td>0.66</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper599</td>
<td>HBC599</td>
<td>18</td>
<td>515/599</td>
<td>54 400</td>
<td>0.43</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>Pepper620</td>
<td>HBC620</td>
<td>6.1</td>
<td>577/620</td>
<td>100 000</td>
<td>0.58</td>
<td>43</td>
<td>106</td>
</tr>
<tr>
<td>SBB-2</td>
<td>SR-DN</td>
<td>1400</td>
<td>579/596</td>
<td>50 200</td>
<td>0.98</td>
<td>75</td>
<td>107</td>
</tr>
<tr>
<td>o-Coral</td>
<td>Gemini-561</td>
<td>73</td>
<td>580/596</td>
<td>141 000</td>
<td>0.58</td>
<td>150</td>
<td>180</td>
</tr>
<tr>
<td>RhoBAST</td>
<td>TMR-DN</td>
<td>15</td>
<td>564/590</td>
<td>96 000</td>
<td>0.57</td>
<td>55</td>
<td>181</td>
</tr>
<tr>
<td>DNB</td>
<td>TMR-DN</td>
<td>350</td>
<td>555/582</td>
<td>47 150</td>
<td>0.9</td>
<td>75</td>
<td>107</td>
</tr>
<tr>
<td>DNB</td>
<td>SR-DN</td>
<td>800</td>
<td>572/591</td>
<td>50 250</td>
<td>0.98</td>
<td>75</td>
<td>107</td>
</tr>
<tr>
<td>BHQ apt (A1)</td>
<td>Cy3-BHQ1</td>
<td>4700</td>
<td>520/565</td>
<td>N/A</td>
<td>N/A</td>
<td>60</td>
<td>182</td>
</tr>
</tbody>
</table>

N/A, not available; $\epsilon$, absorption coefficient; $\Phi$, quantum yield. $E_x/E_m$, excitation/emission wavelength peak value.
RNAs of interest. One fluorogenic RNA applied for this purpose was developed in the Jaffrey Lab in 2011 and was named Spinach. Spinach can activate the fluorescence of a 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) chromophore. Several Spinach derivatives, such as Spinach2, Baby Spinach, and Broccoli, have been further developed with reduced size, improved brightness and folding capability. More recently, imaging of single cellular RNAs has been achieved by using a tandem array of exceptionally bright fluorogenic RNA aptamer/chromophore pairs including Mango II/TO1-B, Pepper/HBC, Riboglow/ATTO590, and Broccoli/Bl. Meanwhile, a wide spectral range ($\lambda_{ex}$, ~380–650 nm; $\lambda_{em}$, ~420–660 nm) of fluorogenic RNA tags is now available for various imaging purposes (Table 1 and Fig. 4).

Instead of using chemical chromophores, fully genetically encoded fluorogenic RNA systems have also been developed based on RNA aptamers that can stabilize and activate FP fluorescence. Here, a bifunctional tDeg peptide, which can promote protein degradation, is fused to the C-terminus of fluorescent proteins. In the presence of an RNA aptamer, the tDeg peptide can bind with the aptamer and inhibit the proteasome-mediated FP degradation. As a result, a fluorescence signal can be generated. By adding ten concatenated RNA tags to a target mRNA, single mRNA molecules can also be imaged in living cells.

Besides functioning as a fusion tag, fluorogenic RNA aptamers have also been used to construct various interesting dynamic RNA nanodevices for cellular target detection and imaging. In these nanodevices, generally, the fluorogenic RNA aptamers are designed to be unfolded initially, resulting in an off-state with a low background fluorescence signal. Once the target is present, a conformational change in the RNA nanodevice is induced to re-fold the fluorogenic RNA aptamers and further activate the fluorescence signal. As a result, these fluorogenic RNA-based nanodevices can be used to detect a wide variety of analytes of interest. In the following sections, we will discuss the design principles and applications of these smart RNA sensors.
3.2 Split fluorogenic RNA aptamer-based nanodevices

One design strategy to regulate the formation of fluorogenic RNA aptamers is based on the split version of these aptamers. By dividing fluorogenic RNA aptamers into two separate fragments, the chromophores cannot bind with the aptamers and thus exhibit minimal background fluorescence signal. The target analyte is normally designed to bind with both fragments of the split aptamer and bring them into proximity. As a result, fluorogenic RNA aptamers can be reassembled to activate the fluorescence of chromophores.

For example, based on a split version of Broccoli, the Fan group has developed an aptamer-initiated fluorescence complementation method for imaging endogenous RNAs in living mammalian cells, including HeLa and human umbilical mesenchymal stem cells (HuMSC). Here, each fragment of the split Broccoli was appended with a sequence complementary to the target mRNA. The target hybridization will then induce the formation of an active DFHBI-binding site, as well as fluorescence activation (Fig. 5a). By replacing the target recognition sequences, these split RNA aptamer-based nanodevices can be easily designed and extended for imaging different endogenous RNA targets. As another example, the Burke group has recently applied a similar split-Broccoli system to monitor intracellular RNA–RNA interactions in E. coli cells. Only when two RNA strands bind with each other, the fluorescence signal of the cognate fluorophore can be activated.

These split fluorogenic RNA aptamers can also be used to improve the sensitivity of RNA-based nanodevices. For example, our group has developed a fluorogenic RNA-based genetically encoded RNA circuit in 2018, termed CHARGE, by combining a catalytic hairpin assembly (CHA) system with split Broccoli (Fig. 5b). CHA is an efficient enzyme-free amplification system based on the target-induced catalytic hybridization of two hairpin structures. By conjugating split Broccoli fragments, respectively, to the terminal of these two hairpins, the target RNA-induced CHA activation will lead to the reassembly of the Broccoli structure and activate the DFHBI fluorescence. One target can catalytically generate tens-to-hundreds of Broccoli, and as a result, the CHARGE circuit can be used to image RNA targets in living E. coli cells with very high sensitivity.

One potential limitation of the CHARGE device is that it cannot be used to track the subcellular location and distribution of the targets. To solve this problem, we have also engineered another split Broccoli-based circuit, which was named IN SItu Genetic Hybridization Amplification Technique (INSIGHT). INSIGHT functions based on a hybridization chain reaction between a pair of split Broccoli-modified hairpins. Once a target RNA is generated inside cells, a cascaded hybridization of these two hairpins will be triggered to generate a chain of Broccoli (Fig. 5c). Because the generated Broccoli chain will directly bind with the target RNA, the INSIGHT system can be used to image the distribution and cellular location of the target analytes in both bacterial and mammalian cells.

3.3 Allosteric fluorogenic RNA aptamer-based nanodevices

Another commonly used design strategy of fluorogenic RNA-based sensors is based on the target-induced allosteric structure change of the RNA aptamers. In this design principle, a target-binding aptamer is directly connected to the fluorogenic RNA aptamer through a transducer RNA module (Fig. 6a). In the absence of the target, both the target-binding aptamer and...
fluorogenic RNA aptamer are unfolded. The binding of the target will then stabilize the transducer and refold the fluorogenic RNA aptamer, turning on the fluorescence. Using Spinach as the reporter, the Jaffrey group developed a type of allosteric fluorogenic RNA sensor in 2012 for imaging the cellular dynamics of adenosine 5'-diphosphate and 5'-adenosyl methionine in E. coli cells.121 Based on a similar design, several allosteric Spinach-based sensors have been further developed for the intracellular imaging of various small molecules and proteins in bacterial cells, including cyclic di-AMP,122 cyclic di-GMP,123 cyclic AMP-GMP,123 streptavidin,124 MCP coat protein,124 etc. In addition to Spinach, other fluorogenic RNA aptamers, such as Broccoli and red Broccoli, have also been applied to construct allosteric sensors for imaging 5-hydroxy-L-tryptophan,125 3,4-dihydroxy-L-phenylalanine,125 S-adenosyl methionine,126 etc. in living bacterial and mammalian cells.

Although a number of allosteric fluorogenic RNA sensors have been developed to image cellular analytes, almost all of these sensors are developed based on a single RNA fluorescent reporter. Considering the variations in the cellular RNA expression and distribution, it is difficult to directly apply these single-color sensors to quantify target cellular concentrations. Our group has developed a ratiometric fluorogenic RNA device to solve this problem.127 Our ratiometric sensor contains two fluorogenic RNA aptamer/chromophore pairs, Broccoli/DFHBI and DNB/SR-DN (Fig. 6a). The Broccoli and DNB aptamers were connected via a three-way junction F30 scaffold. The DNB aptamer was further engineered into an allosteric sensor for imaging 5-hydroxy-L-tryptophan,125 3,4-dihydroxy-L-phenylalanine,125 S-adenosyl methionine,126 etc. in living bacterial and mammalian cells.

3.4 Riboswitch- and ribozyme-based fluorogenic RNA nanodevices

Ribozymes are another type of functional RNA nanodevice enabling catalytic cleavage of RNA substrates at specific positions.131 Due to their easy design, predictable structure, and controllable activity,132 ribozymes have been used as a useful tool for constructing biosensors. By fusing target-binding aptamers with ribozymes, allosteric ribozyme sensors have been...
designed to detect different biomolecules. Based on these allosteric ribozymes and a Broccoli reporter, a type of RNA-based catalytic sensor, named RNA integrators, has been developed for low-abundance metabolite detection in living E. coli cells (Fig. 7b). In the presence of the target analyte, the folding of the target-binding aptamer would activate the ribozyme cleavage, which subsequently triggered the release of an inhibitory Broccoli sequence. As a result, the Broccoli aptamer was reassembled to activate the DFHBI fluorescence. Since each target molecule can bind and induce cleavage of multiple RNA integrators, the fluorescence signal could be amplified.

3.5 Fluorogenic RNA aptamer-based FRET sensors

Another promising sensor design strategy is based on fluorescence resonance energy transfer (FRET). FRET is a non-radiative energy transfer process between an excited fluorophore and a ground-state acceptor. The FRET efficiency is highly dependent on the distance and orientation of the donor and acceptor fluorophores. FRET-based sensors have an inherent sensitivity to the environment and conformation changes, which provides a powerful approach for probing the temporal and spatial variations of target molecules and biological processes.

RNA nanodevices, with defined shape, size, and stoichiometry, have been constructed to precisely assemble different protein and RNA molecules. A very interesting RNA origami structure has been recently used to build a fluorogenic RNA aptamer-based FRET nanodevice in E. coli cells (Fig. 8). Using only a single-stranded RNA, this origami structure can be genetically encoded and used as a scaffold to position two fluorogenic RNA aptamers in different orientations and proximities. After optimizing the distance and relative dipole moment, the FRET signal between a Spinach and a Mango RNA aptamer was used for the detection of target RNAs and small molecules. Upon target binding, the conformation of the RNA origami was altered, which further changed the orientation and distance between Spinach and Mango, leading to changes in the FRET outputs. Further optimized FRET-based RNA nanodevices can be potentially used for the quantitative and rapid imaging of various target analytes.

4. Genetically encoded RNA nanodevices for other cellular functions

4.1 Genetically encoded RNA nanodevices to perform cellular logic operations

Natural biological systems are always present in a complex environment that requires rapid sensing of multiple input signals, a logic analysis, and then an accurate output response. Inspired by the natural gene network, sophisticated DNA-based circuits have been engineered in vitro for information storage, computing, and diagnostics. While these DNA-based circuits are difficult to be used inside cells, genetically encoded RNA nanodevices have been created to perform intracellular logic operations with the goal of regulating cellular functions in a more precise way.

For example, by coupling both theophylline and TPP aptamers to the 5' untranslated region of a mRNA, the Yokobayashi group has previously engineered Boolean AND and NAND logic gates to control gene expression inside E. coli cells. Similarly, the Smolke group has developed another modular
A significant improvement in the programmability of these RNA nanodevices was achieved based on the toehold switches. After an initial demonstration of a four-input AND logic gate function inside *E. coli* cells, the further optimization of RNA sequences has led to the development of a more complex logic network, for example a 12-input logic circuit, i.e., one of the most complicated synthetic logic expression systems realized in living cells. These toehold switch-based nanodevices can also be designed to repress bacterial gene translation with up to four inputs.

Besides toehold switches, the above-mentioned STAR- and CRISPR-based nanodevices have also been validated to perform two- or three-input logic operations inside *E. coli* cells. These and other genetically encoded RNA nanodevices have further expanded the toolbox of programmable computing units for constructing synthetic intracellular circuits and information network.

### 4.2 Genetically encoded RNA nanodevices for structural biology studies

DNA nanotechnology has been used to construct various highly precise nanostructures that can be used as *in vitro* scaffolds to arrange biological molecules in a specific pattern. These rationally designed DNA nanostructures have been further applied to study and regulate the functions and interactions of various target molecules. Interestingly, these DNA nanostructures can now also be genetically encoded through phagemid in bacterial cells. Compared with chemically synthesized DNA molecules, these intracellular DNA nanostructures can be cost-effectively produced, especially when a large amount of DNAs are needed. Even though these genetically encoded DNA nanostructures can be potentially useful in producing scalable nanodevices for *in vitro* applications, the limited adaptability of these phagemid-based expression systems make it hard to directly apply these DNA nanodevices for cellular analysis or regulation.

In contrast, genetically encoded RNA nanostructures are believed to be more applicable because of the single-stranded nature of cellular RNA molecules. In addition, natural RNA–protein interactions and non-Watson–Crick RNA interactions can also facilitate the construction of complex nanostructures inside living cells. The initial attempt of constructing synthetic genetically encoded RNA nanostructures was based on the self-assembly of short RNA modules. After transcription, these short RNA modules can hybridize with each other to construct one-dimensional or two-dimensional RNA structures. These structures have also been used as scaffolds to bind and spatially organize different proteins for the controlled chemical reactions inside bacterial cells.

More recently, the self-folding of a long RNA strand into a designed nanostructure has been achieved both *in vitro* and inside *E. coli* cells. Compared with multicomponent assembly of short RNA modules, these self-folded single-stranded RNA nanostructures can be more rapidly folded and with a higher yield. Indeed, the co-transcriptional folding of these nanostructures has been successfully demonstrated. This is a critical feature for reducing the potential degradation or misfolding of these RNA nanostructures during intracellular applications. These single-stranded RNA nanostructures have already begun to exhibit interesting cellular functions, for example, in the above-mentioned RNA origami-guided fluorogenic RNA FRET system. With further optimized design and characterization strategy, these versatile self-assembled RNA nanostructures can be highly useful in future to spatially arrange different cellular components for structural biology studies and to regulate cellular interactions.

### 4.3 Genetically encoded photo-responsive RNA nanodevices

Using light to control cellular functions is always an attractive approach because of the high spatial and temporal resolution of light. The idea of using light to regulate the function of genetically encoded RNA nanodevices has been proposed for a while. Several light-regulated RNA switches have been indeed demonstrated in *in vitro* based on the specific recognition of RNA aptamers towards a particular photo-induced isomerization state of the chromophore. Unfortunately the intracellular performance of these chromophore-mediated RNA nanodevices has not yet been validated.

Very recently, exciting photo-responsive RNA nanodevices have been engineered by the Mayer and Möglich groups to regulate cellular gene expression in HeLa cells. In these nanodevices, an RNA aptamer, which can specifically recognize a bacterial light-oxygen-voltage photoreceptor (PAL) under blue light, was inserted into the 5′ untranslated region of a gene reporter. After the PAL–RNA conjugation, the gene expression can be sterically inhibited in both bacterial and mammalian cells. This photo-controlled PAL–RNA interaction has been further used to reversibly regulate the cellular functions of micro RNAs and short hairpin RNAs in HEK293 cells.

Our group has also recently demonstrated a genetically encoded RNA aptamer-based photosensitizer system, termed GRAP, for targeted cell regulation in both prokaryotic and eukaryotic cells. These photosensitizers can generate reactive oxygen species (ROS) upon light irradiation and lead to cell structure damage and photodynamic therapy. In this GRAP system, a DNB aptamer was used to selectively bind with a dinitroaniline quencher and separate it from the attached photosensitizer, which can further result in the restoration of the ROS generation (Fig. 9a). Meanwhile, the formation of the DNB aptamer structure could also be controlled by a target RNA of interest. This stimuli-responsive design has been validated in both *E. coli* and HeLa cells using different RNA targets (Fig. 9b). Wavelength-selective photosensitizing was also demonstrated in this GRAP system. As shown by these initial examples, both reversible and irreversible photo-regulated RNA nanodevices can be potentially used to precisely regulate cell functions.
expressing platforms are still in great need for future eukaryotic and in vivo applications of synthetic RNA nanodevices.

Further advancement in programming algorithm is another critical direction to predict and guide the design of these RNA nanodevices. Even though software such as Mfold\textsuperscript{28} and NUPACK\textsuperscript{29} have been successfully developed for the in vitro calculation and predication of the folding and assembly of nucleic acids. The cellular performance of RNA nanodevices in the events of cotranscriptional folding, target small molecule and protein binding, dynamic switching, and gene regulation is still very difficult to simulate. The dynamic nature and versatile interaction modes of these RNA structures have provided much freedom in designing sophisticated nanodevices. However, it also makes it hard to design and characterize these interactions, especially in the presence of other complicated cellular molecules and environments.

On the other hand, once we could computationally or experimentally understand the correlation between RNA sequences and their intracellular structures and dynamics, dramatic information on the cellular functions of natural non-coding RNAs could also result from these advancements. These naturally existing functional RNA molecules can then further inspire new synthetic RNA nanodevices. Indeed, as mentioned above, riboswitches and ribozymes that are discovered from bioinformatics analysis are important functional units now in building synthetic RNA nanodevices.

So far, natural RNA devices and structures are still way more complicated than the synthetic ones. Such complicity may have resulted in the faster kinetics and larger dynamic range of these natural RNA nanodevices. Understanding the underlying mechanism of these precise and dynamic assemblies is important for the design of new RNA scaffolds in structural biology and for the construction of intelligent RNA networks. To better interpret these design mechanisms and increase the speed of developing functional RNA nanodevices, more reliable in vitro systems to mimic intracellular environment, as well as high-throughput platforms for the direct intracellular characterization, are highly demanded. With these further improvements, we believe, in the near future, genetically encoded RNA nanodevices will perform real intelligent intracellular diagnostics and therapeutics in a way, as good as, if not better than, their natural RNA and protein rivals.

5. Conclusions and future perspectives

In this review, we have discussed the current progress and milestones of using genetically encoded RNA-based nanodevices for intracellular gene regulation, fluorescence imaging, and other interesting applications. In this rapidly emerging and cutting-edge research area, numerous new design principles and functions have been shown recently. We hope the examples illustrated in this review will be helpful in inspiring further development of functional genetically encoded RNA nanodevices. In our own opinion, to allow these synthetic RNA devices to compete with, or to exceed, their protein or natural RNA rivals, there are still several important directions that need additional efforts and breakthroughs.

First of all, most of these synthetic RNA nanodevices have only been validated in vitro or in prokaryotic cells, such as E. coli. Some of these studies are even limited in RNase-deficient E. coli strains. Indeed, expressing these RNA devices at a high level with reduced degradation is still a major challenge, especially in eukaryotic cells. Compared with proteins, normal cellular RNA concentrations are believed to be at least one magnitude lower. In nature, some RNA sequences can be partially protected based on different 5' and 3' structures and base modifications.\textsuperscript{172} Inspired by these natural RNA protection mechanisms, synthetic RNA nanodevices may also be similarly secured in the complex cellular environment. For example, a circular RNA expressing system has been designed by the Jaffrey group to significantly reduce the cellular degradation of synthetic RNA nanodevices.\textsuperscript{173} While with the removal of both 5' and 3' ends, RNA devices have to be further carefully optimized in this circular RNA format. Other robust and reliable RNA

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors gratefully acknowledge the UMass Amherst startup grant, NIH R01AI136789, NSF CAREER, and Alfred P. Sloan Research Fellowship to M. You. The authors also thank other members in the You Lab for useful discussion and valuable comments.


131 L. Ma and J. Liu, iScience, 2020, 23, 100815.


