Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. (Pharmaceutical science) Jingye Zhang

Born in: Anhui, China

Oral Examination: .....

Development of Genetically Encoded Near-Infrared Fluorescent Light-up Aptamers for Live-Cell RNA imaging

Refereers: Prof. Dr. Andres Jäschke

Prof. Dr. Michael Knop

Publications and Awards

The presented work was performed from November 2017 to January 2021 under the supervision of Dr. Murat Sunbul and Prof. Dr. Andres Jäschke.

The publications resulted from this work are listed below.

Journal article:

Jingye Zhang, Lu Wang, Andres Jäschke*, and Murat Sunbul*. A Color-Shifting Near-Infrared Fluorescent Aptamer-Fluorophore Module for Live-Cell RNA Imaging. In preparation.

Poster presentation:

Jingye Zhang, Murat Sunbul, Andres Jäschke. Benzopyran-coumarin Fluorophores for Ratiometric RNA Imaging in Living Cells. EMBL Symposium: Seeing is Believing, Heidelberg, Germany, 9-12 Oct 2019.

Awards:

Chinese government scholarship awarded by China Scholarship Council (CSC), grant No. 201606100059, 12 Sep 2016 to 11 Sep 2020.

HBIGS travel grant for EMBL Symposium: Seeing is Believing. Oct 2019.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Dr. Andres Jäschke. I am really grateful for the opportunity you offered me to study and work in your laboratory. You brought me to the aptamer world and widened my horizon in the chemical biology field. Thank you for all the resources, supports, and freedom you provided. All of these have helped me grow as an independent and well-trained researcher.

Additionally, I would like to thank Prof. Dr. Michael Knop and Dr. Jonas Ries for being members of my TAC committee. Your valuable scientific input and constant encouragement during my TAC meetings have advanced my research work and helped me through the tough times. I would also like to thank Dr. Alessia Ruggieri for being a member of my thesis examination committee and finding time to evaluate my work.

Moreover, I would like to express my heartfelt gratitude to Dr. Murat Sunbul. You taught me so much knowledge and all kinds of skills and guided me throughout the thesis. Not only your scientific achievements but also your passion for science and rigorous academic attitude have inspired and encouraged me for my future endeavors.

Next, I would like to acknowledge all the group members for providing an excellent working environment. Thank you for organizing and participating in the cake sessions, game nights, group weekends, breakfasts, and dinners. I had a wonderful time here in this group, and I cherish every minute I spent with this group of crazy friends. I owe special thanks to the past and present members of our ”imaging office”, Marie Georgens, Regina Wirth, Bastian Bühler, Daniel Englert, and Franziska Grün. Thank you for your scientific discussions, useful suggestions, and willingness to help during the difficult course of the doctoral study. My sincere thanks also go to Yaqing Zhang. You provided so many valuable suggestions that helped me accomplish my research and study smoothly.

Importantly, I would like to extend my sincere gratitude to our technical staff, Daniel Wolf, Tobias Timmermann, Heiko Rudy, and Monika Langlotz, for your help with NMR, MS, and FACS experiments, lab equipment maintenance, and IT support. I am extremely grateful to our secretary Viola Keller for being there all around and helping me with all the bureaucratic administrative affairs. I am also thankful to the Chinese Scholarship Council for the Ph.D. scholarship you awarded me for pursuing my doctoral study in Germany.

Besides, I want to thank all the undergraduate and master students who have done internships with me. Thank you, Elena Papalozi, Leon Tydecks, Martina A. Pogány, and Paul Haubenwallner. Thanks for your contributions to my research work. I also learned a lot from you.

Last but not least, I would like to thank my husband and parents. Your love and supports are indispensable. This thesis would not have been possible without the great confidence and constant supports you have shown in me.

Abstract

Visualizing RNA localization and dynamics in living cells provides new insights into RNA biology. Genetically encoded fluorescent light-up RNA aptamers (FLAPs) that can specifically bind and activate the fluorescence of cell-permeable fluorogenic ligands represent emerging and exciting possibilities for real-time live-cell RNA imaging. Although several FLAPs have been developed so far, there is still a strong demand for FLAPs that fluoresce in the near-infrared (NIR) region where cellular autofluorescence and photocytotoxicity are low.

In the first part, we presented the first color-shifting aptamer-fluorophore system that allows simultaneous imaging of both the NIR-fluorescent aptamer-dye complex and the cyan-fluorescent unbound dye. The ratiometric images (NIR/cyan) obtained from this color-shifting module are particularly useful for correcting differences in the probe's cellular uptake, heterogeneous probe distribution, probe instability, and cell morphology variations. This novel system exploited the environmentally sensitive benzopyrylium-coumarin (BC) hybrid fluorophores, which exist in an equilibrium between a spirocyclic cyan-fluorescent form and a zwitterionic NIR-fluorescent form. First, we evolved a 38-nucleotide (nt) RNA aptamer (**BeCA**) that selectively binds the BC zwitterion with nanomolar affinity. Further variations in the BC fluorophore's chemical structure enabled the BC to shift to the cyan spirocyclic form in the unbound state and provided an emission ratio change (cyan/NIR) as high as 15-fold upon aptamer binding. Using the **BeCA-BC** system, we not only imaged **BeCA**-tagged mRNAs fluorescing in the NIR region but also demonstrated its utility in the ratiometric analysis of target RNAs expressed at different levels in single cells using confocal microscopy.

In the second part, we established a fluorescence-activated cell sorting (FACS)-based aptamer selection platform for spirocyclization-based fluorogenic probes. These fluorophores have shown great advantages such as high cell permeability, low background staining, and strong photostability in background-free labeling of various cellular targets in living cells. Unlike the traditional affinity-oriented aptamer selection method SELEX, this new platform offered the possibility of aptamer evolution directed by both brightness and binding affinity, two key parameters of FLAPs. The FACS-based aptamer selection against the fluorogenic and NIR-fluorescent silicon rhodamine (SiR) and the subsequent *in vitro* activity screening successfully revealed three SiR-binding sequence families. Among them, a 103-nt variant L2-1 showed the highest binding affinity (K_D = 25 nM, turn-on = 4.4-fold) and a 53-nt variant L2-2-3 possessed highest fluorescence turn-on (K_D = 192 nM, turn-on = 5.5-fold) towards SiR.

The two NIR FLAPs developed in this thesis provide valuable tools for live-cell RNA imaging and inspire the development of the next-generation of spirocyclization-based FLAPs.

Zusammenfassung

Die Visualisierung von RNA, insbesondere mit Hinblick auf deren Lokalisierung und Dynamik in lebenden Zellen, bietet neue Einblicke in deren biologische Funktionen. Die Möglichkeit Fluoreszente „Light-up" RNA Aptamere (FLAPs), welche fluorogene, zellpermeable Liganden mit hoher Spezifizität und Affinität binden können, genetisch zu codieren, bietet neue Möglichkeiten für die Echtzeit-Bildgebung von RNA in lebenden Zellen. Obwohl bisher mehrere FLAPs entwickelt wurden, gibt es nur wenige, die im Nahen Infrarot Bereich (NIR) fluoreszieren, wo die zelluläre Autofluoreszenz und Photozytotoxizität gering ist.

Im ersten Teil der vorliegenden Arbeit, haben wir das erste farbverschiebende Aptamer-Fluorophor-System entwickelt, welches die gleichzeitige Detektion des NIR-fluoreszierenden Aptamer-Farbstoff-Komplexes und des Cyan-fluoreszierenden ungebundenen Farbstoffs ermöglicht. Die ratiometrischen Bilder (NIR/Cyan), die durch diese Farbverschiebung erhalten wurden, sind besonders nützlich, um Unterschiede der zellulären Aufnahme des Farbstoffes, dessen heterogenen Verteilung und Instabilität sowie Variationen in der Zellmorphologie korrigieren zu können. Dieses neuartige System nutzt die umgebungssensitiven Benzopyrylium-Cumarin-Hybridfluorophore (BC), die im Gleichgewicht zwischen einer spirozyklischen Cyan-fluoreszenzierenden Form und einer zwitterionischen NIR-fluoreszenzierenden Form existieren. Zunächst entwickelten wir einen 38-Nukleotid (nt) langen RNA-Aptamer (**BeCA**), der das BC-Zwitterion selektiv mit nanomolarer Affinität bindet. Variationen in der chemischen Struktur des BC-Fluorophors ermöglichten es, das Gleichgewicht der ungebunden BC-Derivate in Richtung der spirozyklischen Cyan-fluoreszierenden Form zu verschieben. Somit konnte eine bis zu 15-fache Änderung des Emissionsverhältnisses (Cyan/NIR) durch Bindung des **BeCA**s erreicht werden. Das optimierte **BeCA-BC**-System ermöglichte nicht nur die konfokale Fluoreszenzmikroskopie von mRNAs, sondern zeigte auch die Vorteile ratiometrischer Analyse der markierten RNAs in lebenden Bakterien.

Im zweiten Teil etablierten wir eine auf Durchflusszytometrie (FACS) basierende Aptamer-Selektionsplattform für spirozyklisierende, fluorogene Sonden. Diese Fluorophore zeigten große Vorteile, wie eine hohe Zellpermeabilität, eine geringe Hintergrundfärbung und eine starke Photostabilität bei einer hintergrundfreien Markierung verschiedener zellulärer Ziele in lebenden Zellen. Im Gegensatz zur traditionellen affinitätsorientierten Aptamerselektionsmethode SELEX, bietet FACS die Möglichkeit der Aptamerentwicklung, wobei direkt nach der Helligkeit als auch der Bindungsaffinität selektiert wird, welches beides Schlüsselparametern von FLAPs sind. Es konnten drei SiR-Bindungsfamilien durch die FACS-basierte Aptamer-Selektion gegen das fluorogene und NIR-fluoreszierende Silizium-Rhodamin (SiR), gefolgt von einem *in vitro* Aktivitäts-Screening, identifiziert werden. Von diesen drei Familien zeigte eine 103-nt Mutante (L2-1) die höchste Bindungsaffinität (K_D = 25 nM, Floureszenzsteigerung = 4,4-fach) und ein weitere

53-nt Mutante (L2-2-3) generierte die höchste Fluorogenität (K_D = 192 nM, Floureszenzsteigerung = 5,5-fach) durch Bindung des SiR-Fluorophors.
Die beiden in dieser Arbeit entwickelten NIR-FLAPs stellen wertvolle Werkzeuge für die Bildgebung von RNAs in lebenden Zellen dar und inspirieren die Entwicklung der nächsten Generation von FLAPs gegen spirozyklisierende Fluorophore.

Table of contents

1 Introduction

Ribonucleic acid (RNA) is a linear polymeric molecule of nucleotides. Each nucleotide consists of a ribose sugar, a phosphate group, and a nucleobase (Figure 1). The 5′-phosphate group of one nucleotide is linked to the 3′-hydroxyl group of another, forming a covalent phosphodiester bond between two adjacent nucleotides. Since each phosphate group contains a negative charge, the RNA overall is a polyanion. The nucleobases attached to the 1′-position of the ribose sugar are adenine (A), guanine (G), cytosine (C), and uracil (U) for natural RNA. According to their chemical structures, adenine and guanine are classified as purine bases, while cytosine and uracil are classified as pyrimidine bases. Purine and pyrimidine bases form complementary base pairs via hydrogen bonding (A pairs with U, G pairs with C). RNA is generally considered less stable and prone to hydrolysis compared to deoxyribonucleic acid (DNA) due to the difference in the chemical structure. The ribose sugar of RNA possesses a chemically reactive hydroxyl group (-OH) at the 2′-position. Besides, RNA uses the nucleobase uracil instead of thymine in DNA, which lacks a methyl group at the 5′-position.

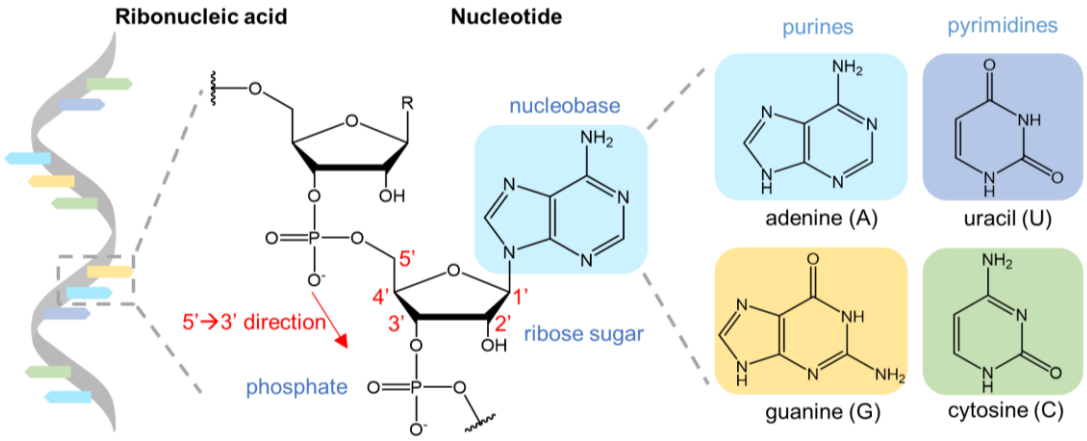


Figure 1. The chemical structures of single-stranded RNA and its building block- nucleotide.

RNA is assembled as a single-stranded polymer chain of nucleotides but can form bulges and helices due to the extensive intramolecular base-pairing within the molecule. The rich diversity and complexity of the three-dimensional structure of RNA are essential for its stability and various functions. For example, the N^6-methyladenosine (m^6A) methylation of RNA by the cellular enzymes METTL3 and METTL14 alters the RNA structure and consequently regulates the RNA stability, localization, transport, and translation[1].

Known as the "central dogma" of molecular biology, RNA plays a vital role in gene expression and regulation (Figure 2)[2]. One important feature of RNA is that it converts the genetic information stored in DNA into proteins. Three types of RNA are majorly involved in this process. Messenger RNA (mRNA) carries the genetic information of protein-coding genes and is transported from the nucleus to the translation site after transcription. Transfer RNA (tRNA) carries amino

acids to the translation site and is assembled and directed by recognizing the three-nucleotide unit (codon) in mRNA. As the last step of translation, ribosomal RNA (rRNA) catalyzes peptide bond formation between amino acids and the growing polypeptide chain in the protein synthesis machinery.

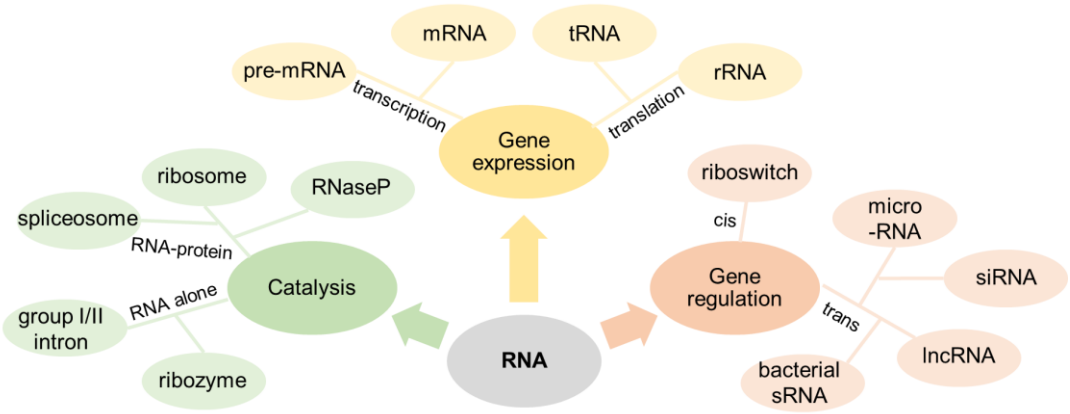


Figure 2. The many roles of RNA in biology (adapted from Ray et al. 2017[2]).

Beyond the canonical role in coding for protein, RNA is becoming more widely recognized for its other functions, including catalysis and gene regulation (Figure 2). Many varieties of RNA are found to be catalytic. For example, as mentioned above, rRNA creates the enzymatic core for protein production in ribosomes. M1 RNA cleaves off the extra sequences and actives precursor-tRNA in ribonuclease Ps[3]. Small nuclear RNA (snRNA) removes the non-protein-coding region from pre-messenger RNA (pre-mRNA) to give mature mRNA in spliceosomes[4]. Besides these RNA-protein complexes, small RNA motifs such as group I/II self-splicing introns are shown to catalyze their own cleavage and ligation reactions[5]. These RNA molecules that carry out specific biochemical reactions similar to protein enzymes are commonly referred to as ribozymes. Besides catalytic properties, many other types of RNA have shown complex regulatory roles in cells. In prokaryotic cells, riboswitches that can change conformation upon the binding of small-molecule ligands were identified[6]. They act as RNA sensors in response to environmental or metabolic changes and affect the expression of downstream RNAs accordingly. Other bacterial small RNAs (sRNAs) modulate protein activity and mRNA stability through specific binding/base-pairing[6]. In eukaryotes, RNA regulators such as small interfering RNA (siRNA) and microRNA (miRNA) were discovered[7]. siRNA is a group of exogenous double-stranded short non-coding RNA molecules, while miRNA is a class of endogenous single-stranded small non-coding RNA molecules. They form RNA-induced silencing complexes (RISCs) with an endo-ribonuclease called Dicer and many other associated proteins inside cells. The RISC binds to the complementary mRNA target using the siRNA/miRNA as a template and silences the gene by cleaving the mRNA. Besides these small non-coding RNAs, long non-coding RNAs (lncRNAs) were also found in mammals and other

eukaryotes. They have various transcriptional and post-transcriptional functions and direct epigenetic regulation[7].
RNA is involved in many biological processes. Visualizing the subcellular location and distribution of RNA can help to understand the rich dynamics of RNA and reveal the mechanism behind its intricate functions. Traditional biochemical assay for RNA analysis could only provide averaging information from billions of molecules fractionated at certain time points, thereby losing cell-specific information. The recent development of fluorescence microscopy and RNA labeling methods opens exciting possibilities for intracellular RNA imaging, providing the advantages of RNA visualization in an intact cell at a high spatial resolution.

1.1 RNA imaging *in situ*

Fluorescence *in situ* hybridization (FISH) is the most commonly used RNA imaging technique for fixed samples. FISH employs the fluorophore-labeled sequence-specific oligonucleotides to reveal RNA transcripts upon complementary hybridization. The concept of imaging nucleic acid via hybridization was first reported in 1969 by Pardue *et al*.[8] They used the radioactive test DNA probe that specifically binds a target DNA via hybridization. This method allowed the visualization of the cellular location of the specific DNA fractions under autoradiograph. Since then, FISH has been applied to identify specific nucleic acid sequences as well as reveal the sites for RNA processing, transportation, and localization.
In 1998, Femino *et al*. made the breakthrough in single-molecule FISH (smFISH)[9]. Multiple oligonucleotides that bind along a target RNA were labeled with five fluorophores per molecule (Figure 3A). Using those probes, the authors successfully achieved quantitative β-actin mRNA imaging by creating diffraction-limited spots. Instead of long oligonucleotides (~ 50-nt) with multiple fluorophores on each, mono-fluorophore labeled short oligonucleotides (~ 20-nt) were generated to reduce the background signal from the mismatched or unbound probe (Figure 3A).
FISH with sequential tethered and intertwined oligonucleotides complexes (FISH-STICs) were designed to amplify the signal further and reduce the preparing cost (Figure 3B)[10]. FISH-STICs consist of a primary probe, an amplifier, and detector probes. The primary probe contains a recognition sequence and a readout sequence. The recognition sequence binds to the target RNA through the sequence-specific hybridization. The amplifier is then added and binds to the primary probe through the readout sequence. Multiple detector probes (i.e., fluorophore-labeled oligonucleotides) hybridize to the amplifier, thereby labeling the target RNA. This system can be easily applied to image different RNAs of interest (ROIs) by simply changing the sequence-specific region of the primary probe. Hence, the preparation procedure and cost are significantly reduced. Other amplification methods include the hybridization chain reaction (HCR)-FISH[11] and padlock-FISH[12]. HCR-FISH uses a primary probe to bind the

target RNA in a sequence-specific manner (Figure 3C)[11]. A pair of fluorophore-labeled nucleotides that form metastable hairpins are then added. The hairpins bind the readout sequence and start the chain reaction. The signal is amplified by forming a long polymer labeled with multiple fluorophores. Alternatively, the padlock probe is a long oligonucleotide whose ends bind to the target sequence at adjacent positions, allowing the circularization by ligation (Figure 3D)[12]. A universal amplification (e.g., rolling circle reaction) is then applied to amplify the signal. This method is particularly useful for short RNAs (0.5-1 kb) since their lengths limit the simultaneous binding of multiple detectors.
The recent advance of smFISH, including multiplexed error-robust (mer) FISH and sequential (seq) FISH, allows RNA multiplexing up to 10000 species within one cell (Figure 3E)[13-14]. First, an oligonucleotide mixture (up to one million) is prepared by on-chip synthesis. Each probe contains one recognition sequence and two readout sequences. The probes hybridize to the target RNAs and label them with multiple readout sequences. The detector probes are then added through sequential rounds of hybridization. In theory, 2^n-1 RNA species can be differentiated (n: number of readout sequences) after decoding.

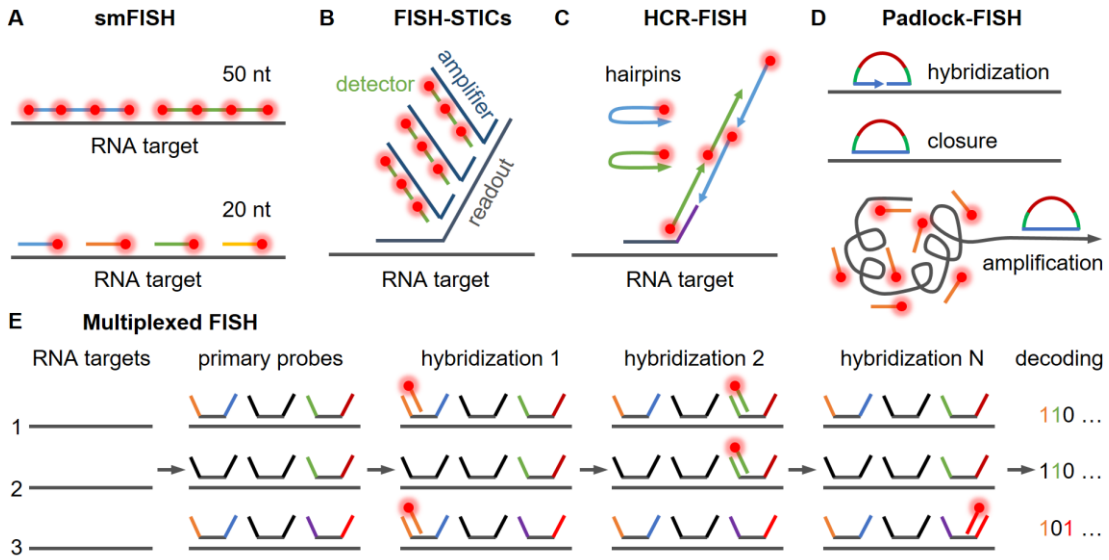


Figure 3. FISH detection of RNA in fixed cells (adapted from Pichon et al. 2018[15]). (A) smFISH. RNA targets are labeled with long oligonucleotides (~50 nt) that carry multiple fluorophores (red) or short oligonucleotides (~20 nt) that carry a single fluorophore (red). (B) FISH-STICs. RNA target is hybridized with the recognition sequence of a primary oligonucleotide, while the readout region of the primary oligonucleotide is hybridized with several amplifier oligonucleotides labeled with multiple fluorescent detector oligonucleotides. (C) HCR-FISH. Two hairpin structured oligonucleotides bind along the readout sequence of a primary oligonucleotide, forming a long fluorophore-labeled polymer. (D) Padlock-FISH. A padlock probe binds to a target RNA, forming a circular structure with a nicked site, which is further repaired by a ligase. The fluorescent signal is then amplified using a universal amplification method such as rolling circle reaction. (E) Multiplexed FISH. Primary probes contain a recognition sequence flanking by two readout sequences. The cellular RNA targets are detected by successive rounds of hybridization and identified by decoding.

FISH and its variants open up scientists' understanding of RNA biology. Compared to the traditional biochemical assay, FISH allows a direct RNA visualization in a cellular environment. Besides, it is ideally suited for low-abundant RNAs due to the capability of signal amplification. The recently developed multiplexing techniques combined with advanced probe design algorithms and image analysis tools allowed the detection of massive RNA species simultaneously and showed the potential of imaging-based transcriptomes. However, besides the tedious and complicated workflow, which can cause the potential artificial error, one big limitation of FISH-based methods is that they provide only a static picture at the time of fixation. In contrast, live-cell imaging methods promise real-time RNA dynamics with finer temporal and spatial resolution and provide powerful new analytical possibilities.

1.2 Live-cell RNA imaging

Live-cell RNA imaging is the fluorescence imaging of a target RNA that is specifically labeled with fluorescent probes. To date, the reported live-cell RNA imaging methods can be classified into three broad categories: (1) hybridization probes; (2) RNA-binding fluorescent proteins; (3) fluorophore-aptamer systems. The applications of these imaging tools have shown successful visualization of various RNA transcripts in living cells and yielded many new insights into RNA biology.

1.2.1 Hybridization probes

Similar to FISH, hybridization probes are fluorophore-labeled sequence-specific oligonucleotides. The fluorescence of the labeling fluorophore is quenched by either contact-mediated quenching or interacting with another fluorophore via fluorescence resonance energy transfer (FRET). They are invasively delivered into cells and bind to the endogenous target RNA. The fluorescent properties change upon hybridization, thus enabling the specific detection of the target RNA. The first-generation hybridization probes are competitive probes (Figure 4A)[16]. A fluorophore and a quencher are attached to the 5′ end and 3′ end of two complementary strands, respectively. The target-binding strand's fluorescence remains quenched due to the close interaction between the fluorophore and the quencher. However, the competitive binding of the target RNA to the fluorophore-possessing strand releases the quencher-possessing strand, thereby inducing strong fluorescence. Notably, the labeling kinetics can speed up by designing a longer target-binding strand with an overhang, which quickly binds to the target RNA and accelerates the strand displacement[17]. Instead of using the quencher, side-by-side probes are pairs of fluorophore-labeled oligonucleotides, which bind to the target RNA in close proximity (Figure 4B)[18]. The 3′ end and 5′ end fluorophores function as a FRET pair and lead to the FRET detection of the target RNA. Since the FRET signal only exhibits when both detector probes bind to the target, the signal specificity is thus significantly increased. Molecular

beacons are single-stranded oligonucleotides that bear a fluorophore and a quencher at either end (Figure 4C)[19-20]. The oligonucleotide forms an internally quenched hairpin configuration by bringing the fluorophore and quencher to the close proximity. Specific binding between the loop region and the target RNA leads to the separation of the fluorophore and quencher, forming a fluorescent probe-target hybrid. Molecular beacons have been used to image various RNAs in several biological contexts[19-20].

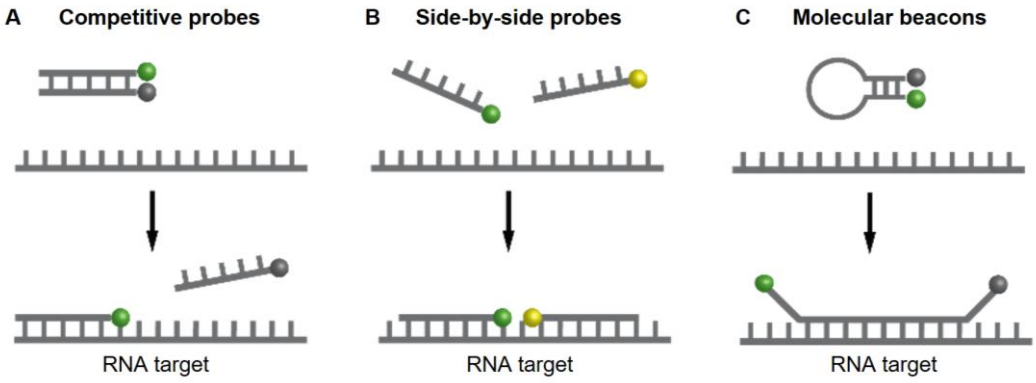


Figure 4. Hybridization probes for RNA detection (adapted from Tyagi et al. 2009[19]). (A) Competitive probes. A double-strand oligonucleotide is labeled with a fluorophore (green) and a quencher (gray). The target RNA binds to the fluorophore-labeled strand by replacing the quencher strand, resulting in the labeling of target RNA. (B) Side-by-side probes. A donor-fluorophore (green) labeled oligonucleotide and an acceptor-fluorophore (yellow) labeled oligonucleotide bind to target RNA at adjacent positions, allowing FRET detection of the target RNA. (C) Molecular beacons. The internally quenched molecular beacon hybridizes to the RNA target, opening up the hairpin structure. The fluorescence is restored due to the separation between the fluorophore (green) and quencher (gray).

The main limitation of these hybridization probes is that they are hydrophilic molecules and cannot move freely across the cell membrane. Invasive probe delivery techniques such as microinjection[21], electroporation[22], lipofaction[23], pore-forming agents[24], and cell-penetrating peptides[24] are needed. Once delivered into cells, these probes quickly accumulate in the nucleus, which is good for labeling nascent RNAs in the nucleus but not cytoplasmic RNAs[25]. To maintain a high cytoplasmic probe concentration, either a cytoplasm-directed moiety (e.g., streptavidin, quantum dot, tRNA) or a slow delivery speed is used[22]. The other limitation is the relatively short half-life (< 30 min) of these oligonucleotides with the natural backbone in cells unless modified nucleotides are used[21, 26].

1.2.2 RNA-binding fluorescent proteins

RNA-binding fluorescent protein systems are the current gold standards for live-cell RNA imaging. A fluorescent protein (FP) is fused to an RNA-binding protein (RBP), which binds to the cognate RNA motif with great affinity and specificity (Figure 5A). Multiple repeats of the RNA motif are genetically fused to an ROI, and the engineered RNA is expressed in cells along with the FP-RBP. The

subsequent binding of the FP-RBP to the RNA motif concentrates the fluorescence, thereby allowing the detection of the ROI. The RBP and the RNA motif should be exogenously derived and should not be found in the target cell's genomic protein/RNA to avoid cross-reactions. MS2 system is the most commonly used FP-RBP tagging system. The coat protein of the bacteriophage MS2 (MCP) binds to a unique stem-loop structure MBS (MS2 binding site) with a strong affinity (dissociation constant [K_D]= 5 nM)[27]. By tagging the FP to MCP, the MS2 system has been used for imaging various RNAs in the cytoplasm in diverse biological contexts. Other similar systems include a 22-amino-acid bacteriophage peptide λ_N and its tight RNA binder boxB (K_D = 22 nM)[28], and the coat protein of phage PP7 (PCP) and its binding RNA motif PBS (PP7 binding site; K_D = 1 nM)[29].

Further approaches such as split-FP were developed to reduce the background signal from the unbound FP[30-32]. The FP is split into two non-fluorescent fragments, to which a pair of RBP tags are attached respectively. Only when the RBP tags are brought into proximity by binding to the RNA motif can the split-FP assemble into a correctly folded protein that is fluorescent. For example, Rackham *et al*. fused the split Venus protein to two RBPs (MCP and a zipcode–binding protein; Figure 5B)[30]. When the split-FPs are brought together through recognizing the MBS and β-actin zip code G quartet, the fluorescent Venus is formed. Instead of two independent RBPs, Valencia-Burton *et al*. dissected eukaryotic initiation factor 4A (eIF4A) into two fragments RBPF1 and RBPF2 (Figure 5C)[31]. The split enhanced green fluorescent protein (EGFP) is reconstructed and recovers fluorescence when RBPF1 and RBPF2 bind to either side of the RNA motif. Ozawa *et al*. took an alternative way by introducing PUMILIO1, a unique RBP that binds to RNA in a sequence-specific manner instead of recognizing the secondary structure (Figure 5D)[32]. The PUMILIO1 can be engineered to target different RNA sequences. The split-FP fused PUMILIO1s recognize two RNA stretches at adjacent positions of the endogenous target RNA, bringing the two halves of FP together. The fluorescence of FP is restored afterward.

Clustered regularly-interspaced short palindromic repeats (CRISPRs) were discovered as adaptive immune systems in bacteria and archaea[33]. Microbes use CRISPR RNAs and CRISPR-associated nucleases (Cas) to defend the invading genetic molecules. The CRISPR/Cas systems have been modified for genome editing, modulating gene expression, and imaging genomic loci in eukaryotic cells. Nelles *et al.* first reported the utility of CRISPR/Cas to image RNA (Figure 5E)[34]. The RNA-targeting FP-fused endonuclease-deficient Cas protein (RCas9) binds to an ROI in a sequence-specific manner in the presence of a guide RNA (gRNA). After being transported to the cytoplasm, the accumulation of the ROI as RNA granules was observed. Recently a more user-friendly system that uses the RNA-guided and RNA-targeting inactive Cas13 ribonuclease (dCas13) for RNA imaging was reported[35]. The authors showed the robust real-time imaging and tracking of endogenous NEAT1, SatIII, MUC4, and GCN4 RNAs in living cells using FP-dCas13.

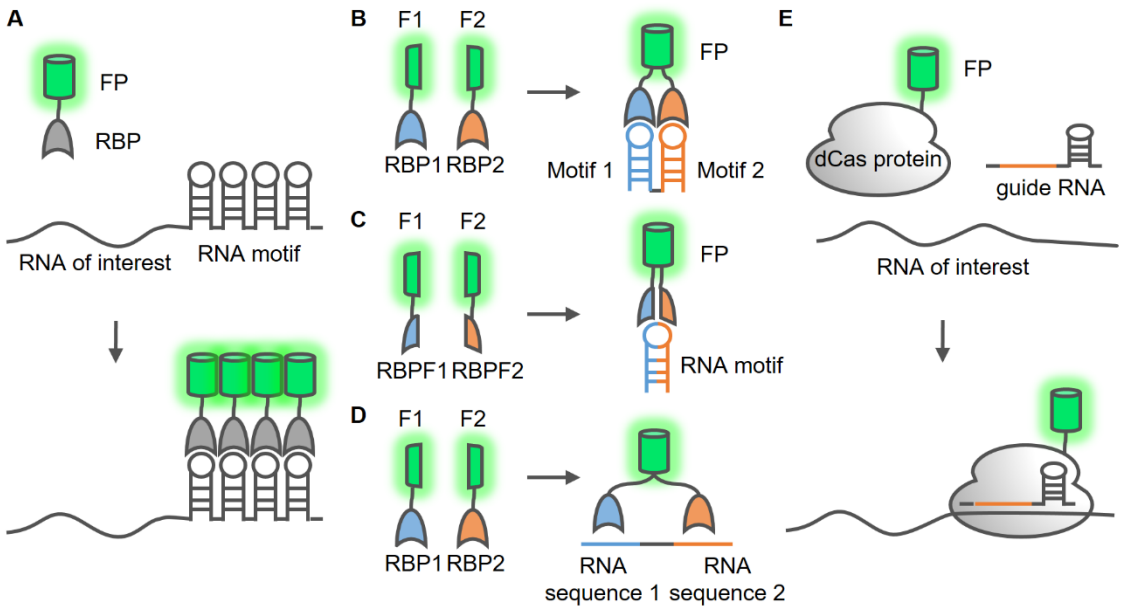


Figure 5. RNA-binding fluorescent proteins for live-cell RNA imaging. (A) Multiple FP-RBPs bind to an array of RNA motifs, allowing the specific detection of an ROI. (B) FP fragment 1 (F1) and FP fragment 2 (F2) are fused to two RNA-binding proteins (RBP1 and RBP2). Each RBP recognizes its cognate RNA motif, and the fluorescence is restored due to the reconstruction of the split-FP. (C) F1 and F2 are fused to the split-RBP (RBPF1 and RBPF2). Each RBPF recognizes half of the RNA motif, mediating the assembly of the functional FP. (D) F1 and F2 are fused to two RBPs, which are engineered to bind to adjacent endogenous RNA sequences. The binding of these fusions to the target RNA reconstitutes FP fluorescence. (E) The programmable RBP (dCas) is fused with an FP and binds the target RNA in the presence of a gRNA.

The RNA-binding fluorescent proteins are great labeling tools for imaging target RNAs in living cells. Nevertheless, FP-RBPs are always fluorescent, making it hard to distinguish between the unbound and bound FP-RBP. Researchers can either express a low number of FP-RBPs or use an array of FP-RBP binding motifs to circumvent this problem. Split FP-RBP approach that the RBP only becomes fluorescent after specific binding is also an effective way. Besides, the real-time imaging of RNA dynamics is difficult by using FP-RBPs since the time required for FP maturation leads to a signal delay. Moreover, the RNA tag size, the insertion point, and the add-on weight from FP-RBPs might alter the function, localization, and stability of the target RNA[19]. Careful control experiments should be designed and performed to validate the findings.

1.2.3 Fluorescent light-up RNA aptamers

For the past decades, fluorescent light-up RNA aptamers (FLAPs) have provided robust and versatile tools for live-cell RNA imaging. The FLAP system consists of two key components: an RNA aptamer and a fluorogenic ligand.
RNA aptamers are short single-stranded RNA oligonucleotides (10-100 nt) that specifically bind target molecules[36]. They were first reported by Szostak and Gold in 1990 as RNA molecules selectively interacting with organic dyes[37] and T4

DNA polymerase, respectively[38]. Aptamers are usually identified by performing *in vitro* selection from a large random sequence pool for specific target molecules, namely systematic evolution of ligands by exponential enrichment (SELEX). Like proteins, aptamers fold into unique tertiary structures, possessing a combination of stems, loops, bulges, pseudoknots, and G-quadruplexes. This allows recognition of target molecules specifically and non-covalently through π - π stacking, hydrogen bonding, electrostatic interactions, and Van der Waals forces. Aptamers are recognized as antibody mimics and have been developed to detect entities ranging from metal ions, small molecules, proteins to whole cells.
Fluorogenic ligands are a group of small-molecule fluorophores, which can activate their fluorescence signal upon environmental changes. They are "dark" in aqueous solutions due to different quenching mechanisms while interacting with their targets (e.g., proteins, aptamers) significantly induces their fluorescence. The specific requirements for small-molecules as fluorogenic ligands for FLAPs will be discussed in the next chapter (chapter 1.3).
FLAPs that utilize cell-permeable fluorogenic ligands have been successfully applied for live-cell RNA imaging (Figure 6). The aptamer tags are genetically engineered to an ROI and are transcribed in cells using the cellular transcription machinery. After supplementing the cell-permeable fluorogenic ligands, the aptamers specifically recognize the ligands and activate their fluorescence. Thus, the aptamer-tagged ROIs are illuminated with high signal-to-background ratios.

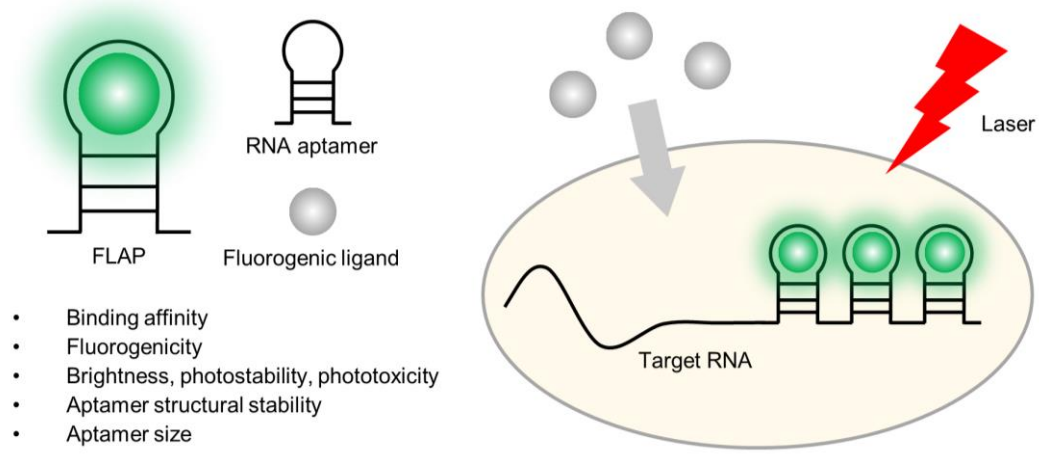


Figure 6. Fluorescent light-up RNA aptamers for RNA imaging in living cells. RNA aptamers are genetically fused to a target RNA and transcribed by the cellular machinery. The fluorogenic ligands pass the cell membrane and become fluorescent upon aptamer binding.

The ultimate goal of RNA imaging is the real-time tracking of endogenous RNAs at a single-molecule level with minimized disturbances. To achieve this, several important criteria for developing FLAPs have to apply. (1) Binding affinity. The specific interaction between aptamers and ligands is governed in a non-covalent manner. A high binding affinity, ideally down to a low-nanomolar or picomolar range, is always preferred considering the low concentration of endogenous RNAs. Besides, the binding kinetics needs to be regulated for different imaging

techniques. A slow dissociation rate of the ligand from the aptamer that allows a long fluorescence "on" time is favored by conventional fluorescence microscopy. Meanwhile, the aptamer-fluorophore pair with fast binding and dissociation rates that offer intermittent fluorescence emission facilitates imaging by single-molecule localization microscopy (SMLM). (2) Fluorogenicity. RNA imaging using FLAPs gains high contrast by illuminating only the specific signal from the aptamer-fluorophore complex but not the unbound probe. Therefore, the fluorescence of the ligand should be fully quenched in the unbound state. A remarkable fluorescence turn-on is only achieved through the interaction between the aptamer and the fluorogenic ligand. Unspecific intracellular staining of the ligand should be suppressed to reduce the background noise. (3) Phototoxicity, brightness, and photostability. Phototoxicity is mainly generated from the utilization of strong laser power or the reactive oxygen species (ROS) generating fluorophores. It should be minimized to reduce the potential disturbance to normal cellular status. Therefore, FLAPs with high brightness are highly preferred to enable sensitive detection under low laser illumination. Besides, FLAPs with high photostability are also recommended to achieve real-time and long-term RNA tracking. (4) Aptamer structural stability. The correct folding of aptamers is a prerequisite for FLAPs functioning in a cellular environment under physiological conditions. Therefore, aptamer structural stability should be carefully assessed in the aspects of thermal stability, cation dependence, existence of G-quadruplex, resistance to enzyme degradation, as well as the insertion point of the aptamer tag. (5) Aptamer size. Tandem repeats of aptamer tags are introduced to an ROI to increase the signal-to-noise ratio. The aptamer size should be kept as minimal as possible to minimize the potential disturbance of the normal function and location of the ROI.

1.2.3.1 Malachite green-binding aptamer

The first selective evolution of a fluorescent light-up aptamer MGA (malachite green-binding aptamer) that binds to a small membrane-permeable triphenylmethane dye, malachite green, was reported by Babendure *et al*. in 2003 (Figure 7)[39]. Malachite green has an extremely low quantum yield in the solution due to vibrational motion. MGA stabilizes the malachite green in a planar, more fluorescent conformation with a high affinity (K_D = 117 nM), inducing a fluorescence enhancement of 2360-fold. The remarkable fluorescence turn-on opens the possibility for RNA imaging by tagging tandem repeats of the aptamer to ROIs. However, the malachite green is found to produce singlet oxygen upon irradiation and therefore considered cytotoxic and not ideal for living cells[40].

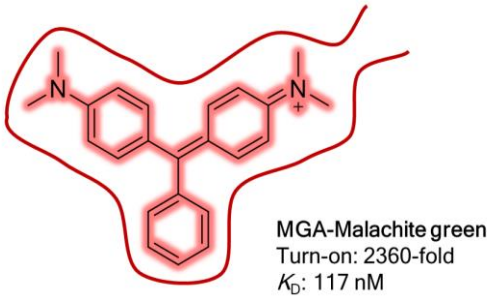


Figure 7. Malachite green-binding aptamer (MGA; red curve) with its bound ligand, malachite green.

1.2.3.2 Spinach and its improved versions

In 2013, Paige *et al*. published the fundamental paper in the FLAP field for live-cell imaging[41]. They exploited 4-hydroxybenzlidene imidazolinone (HBI), a GFP mimic small-molecule dye that only becomes fluorescent after rigidification (Figure 8). RNA aptamers that form complexes with a series of HBI derivatives with emission spanning the whole visible spectrum were isolated after ten rounds of SELEX. Among them, an aptamer termed Spinach (binding core: 80-nt) selectively binds 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) and actives its fluorescence up to 2000-fold with a K_D of 537 nM. This RNA-fluorophore complex was used for 5S rRNA imaging in HEK293 T cells. Later, Spinach2 with improved thermal stability was generated by systematic mutagenesis and exhibited increased intracellular fluorescence when imaging the dynamics of the CGG repeat-containing RNA aggregates[42]. DFHBI-1T, derived from the substitution of a trifluoromethyl group on the imidazolinone ring of DFHBI, exhibited lower background fluorescence and improved excitation and emission wavelengths that are more compatible with standard microscope settings. A brighter, shorter (46-nt), and less magnesium-dependent aptamer dubbed Broccoli was optimized for DFHBI-1T using fluorescence-activated bacteria sorting. The authors successfully imaged 5S rRNA without using a stabilizing tRNA scaffold, which is impossible with Spinach[43]. Moreover, Li *et al*. recently reported a benzimidazole substituted DFHBI derivative, referred to as BI[44]. The Broccoli-BI complex showed improved photostability owing to the impaired photoisomerization of BI and increased thermal stability since BI prevents the thermal unfolding of Broccoli. To produce red-shifted FLAPs, another aptamer Corn (binding core: 28-nt), which binds a red fluorescent protein (RFP) mimic, 3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime (DFHO), was developed[45]. The improved thermal stability and photostability of Corn relative to Spinach and Broccoli enabled the imaging of RNA polymerase III-dependent transcription. Recently, Steinmetzger *et al*. reported a Chili aptamer, which displays a large Stokes shift using an HBI derivative $DMHBO^+$. Interacting with Atto 590, the Chili- $DMHBO^+$ system enabled RNA analysis using the FRET signal[46].

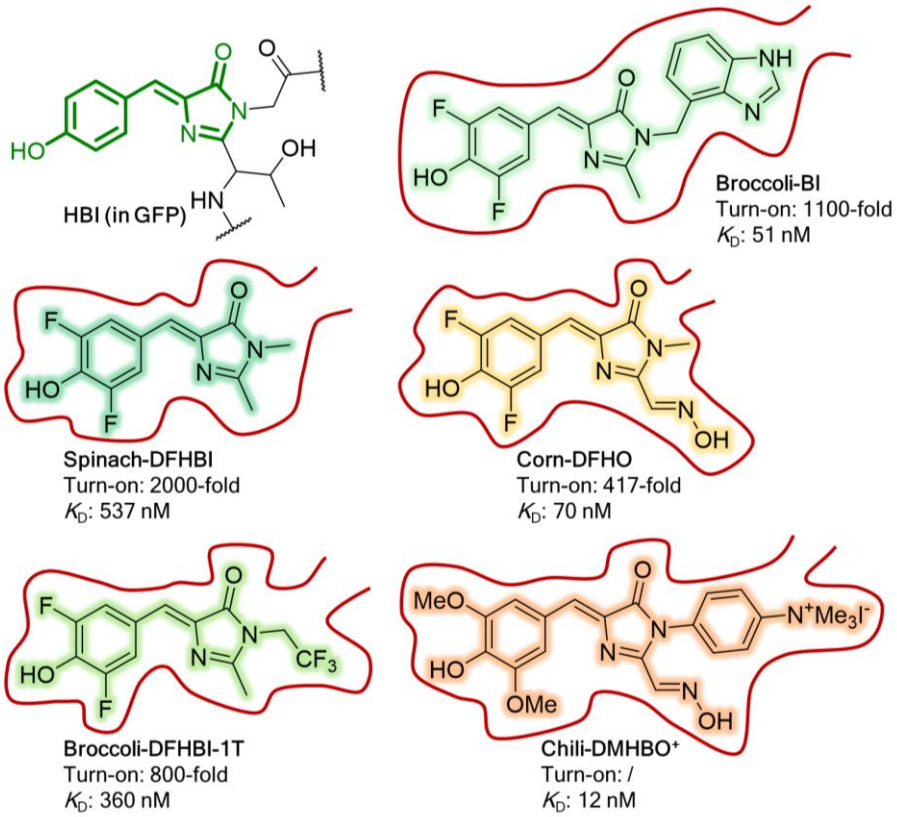


Figure 8. 4-Hydroxybenzylidene imidazolinone (HBI)-binding aptamers (red curve)-Spinach, Broccoli, Corn, Chili, with their bound ligands, DFHBI, DFHBI-1T, BI, DFHO, and DMHBO⁺, respectively.

1.2.3.3 Mango and its improved versions

Instead of HBI-based fluorophores, RNA Mango (binding core: 23-nt) was selected and evolved for a palette of thiazole orange (TO) derivatives (Figure 9)[47]. RNA Mango complexes with TO1-Biotin with an ultra-high binding affinity (K_D = 3.2 nM) and a fluorescent enhancement up to 1100-fold. Due to the attenuated autofluorescence in the near-infrared (NIR) region (650 - 900 nm), a red-shifted TO derivative, TO3-Biotin, was further developed and tested with Mango. It binds to Mango tightly with a K_D value of 6-8 nM but possesses a considerably lower extinction coefficient. By tagging Mango aptamer to bacterial 6S RNA, the authors showed this system's capability to label the freshly transcribed and purified RNAs with TO1-Biotin fluorescently. The same research group in 2018 reported the reselection of Mango, generating three new high-affinity fluorogenic Mangos using a microfluidic-based selection[48]. Mango II, III, and IV retained the high binding affinity (K_D = 0.7, 5.6, 11.1 nM, respectively) but are brighter (1.5-, 4-, and 3-fold higher in brightness, respectively) than Mango I. The TO3-Biotin also showed better fluorescence performance relative to Mango I when incubated with Mango II and IV. The authors further demonstrated the ability of the new Mangos for imaging the subcellular localization of 5S rRNA, snRNA, and scaRNA (a small Cajal-body-specific RNA) in both fixed and living mammalian

cells. More recently, single RNA molecules were revealed with high contrast by Mango II arrays in mammalian cells[49].

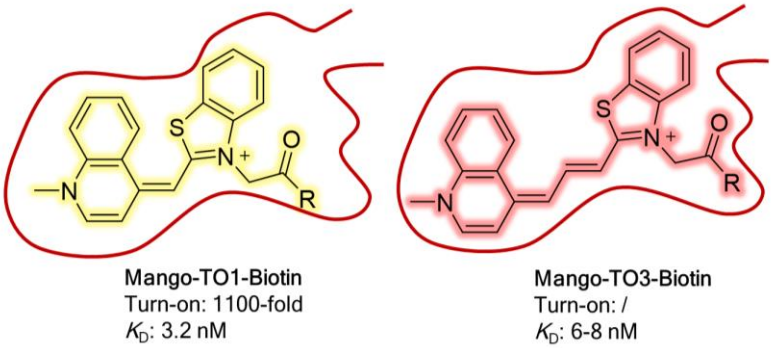


Figure 9. Thiazole orange (TO)-binding aptamer Mango (red curve) with its bound ligands, TO1-Biotin and TO3-Biotin, respectively.

1.2.3.4 Pepper

Analogously, the recently published aptamer Pepper (binding core: 43-nt) was generated to bind a series of (4-((2-hydroxyethyl)(methyl)amino)-benzylidene)-cyanophenylacetonitrile (HBC), activating the fluorescence by constraining the intramolecular motions of these probes (Figure 10)[50]. Pepper-HBC complexes showed remarkable fluorescence turn-ons (~ thousands-fold), high binding affinities (K_Ds in the low-nanomolar range), good thermal stabilities (T_m > 45 °C), and broad emission spectra spanning from cyan to red. Unlike Spinach[51-52], Corn[53], and Mango[54], which contain a G-quadruplex that is not stable/present in mammalian cells[55], Pepper folds independently of potassium and demonstrated good structural stability in mammalian cells, showing an improved intracellular signal relative to Broccoli and Corn (9- and 11-fold, respectively, in terms of Pepper530). Peppers were used to image various RNA species, including 5S rRNA, snRNA, and splicing RNA, without the need for a scaffold RNA. By tagging Pepper to a blue fluorescent protein (BFP) mRNA, Pepper530 revealed the translation kinetics of BFP mRNA and the expression heterogeneity at single-cell level with minimal perturbations. Fusing Pepper to the chimeric sgRNA, genomic loci were imaged with HBC620 that colocalized well with FP fused dCas in a CRISPR display. In light of the superior brightness and photostability, Pepper620 further showed its capability in super-resolution imaging of a BFP-histone mRNA using the structured illumination microscopy (SIM).

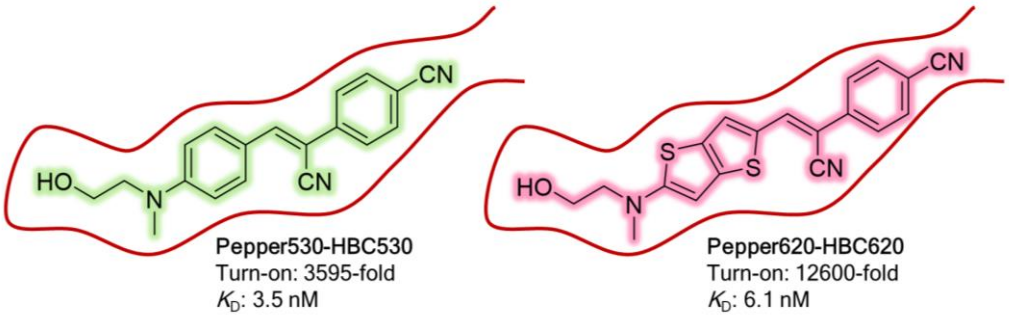


Figure 10. 4-((2-Hydroxyethyl)(methyl)amino)-benzylidene)-cyanophenylacetonitrile (HBC)-binding aptamer Pepper (red curve) with its bound ligands, HBC530 and HBC620, respectively.

1.2.3.5 Rhodamine-binding aptamer and its improved versions

Rhodamine fluorophores have been widely used for bioimaging due to their high quantum yields, excellent photostabilities, and good cell permeabilities. In 2013, our group published the first FLAP system for RNA imaging based on rhodamine fluorophores (Figure 11)[56]. SRB2 is a 54-nt sulforhodamine B-binding aptamer initially reported by the Szostak group in 1998[57]. Intramolecular quenchers were introduced to sulforhodamine B to promote the formation of non-fluorescent complexes in aqueous solutions. After screening a set of quenching moieties, they identified dinitroaniline (DN) as an efficient contact quencher for sulforhodamine B, generating a fluorogenic probe SR-DN. The binding of sulforhodamine B to SRB2 released the DN from the non-fluorescent complex. A 105-fold fluorescence enhancement was observed with a moderate K_D value of 1.4 μM. The authors demonstrated the capability of this system for real-time *in vitro* transcription monitoring and proof of principle RNA imaging experiment in living bacteria. Further study showed that this SRB2 aptamer has the ability to promiscuously bind to various dyes, which possess similar scaffold structure as sulforhodamine B[58]. Among them, 5'-carboxytetramethylrhodamine (TMR) conjugated with a DN quencher was revealed as a superior fluorogenic ligand for SRB2. TMR-DN showed a greatly improved binding affinity (K_D = 35 nM) with a moderate turn-on (17-fold) but much brighter fluorescence signal inside cells than SR-DN. This system was applied to image 5S rRNA and CFP mRNA in mammalian cells. Using TMR-DN as the ligand, RhoBAST, which was evolved from an SRB2 doped library, exhibited an ultra-high binding affinity (K_D = 15 nM), an increased fluorescence turn-on (26-fold), excellent thermal stability (T_m = 79 °C), and strong photostability. Additionally, RhoBAST showed fast association and dissociation kinetics towards TMR-DN, significantly (> 3 orders of magnitude) different from the other conventional FLAPs. The "blinking" events generated from the fast ligand exchange enabled the visualization of subcellular structures of different RNAs beyond the diffraction limit using SMLM[59].

Instead of a dye-quencher conjugate, Gemini-561 is a self-quenching dimer of sulforhodamine B (Figure 11)[60]. A dimerized SRB2 derived aptamer termed as o-

Coral was generated for Gemini-561 through an aptamer selection by combining SELEX and microfluidic-assisted *in vitro* compartmentalization (μIVC). Gemini-561 improved its performance upon aptamer binding, demonstrating a 13-fold fluorescence turn-on with a K_D of 73 nM. Compared to monomeric SRB2-sulforhodamine B and other reported FLAPs (including Broccoli, Corn, Mango), o-Coral-Gemini-561 showed improved brightness, photostability, and binding affinity. The authors realized the fluorescence detection of polymerase II-expressed EGFP mRNA and polymerase III-expressed 5S rRNA with only one copy of o-Coral in living mammalian cells.

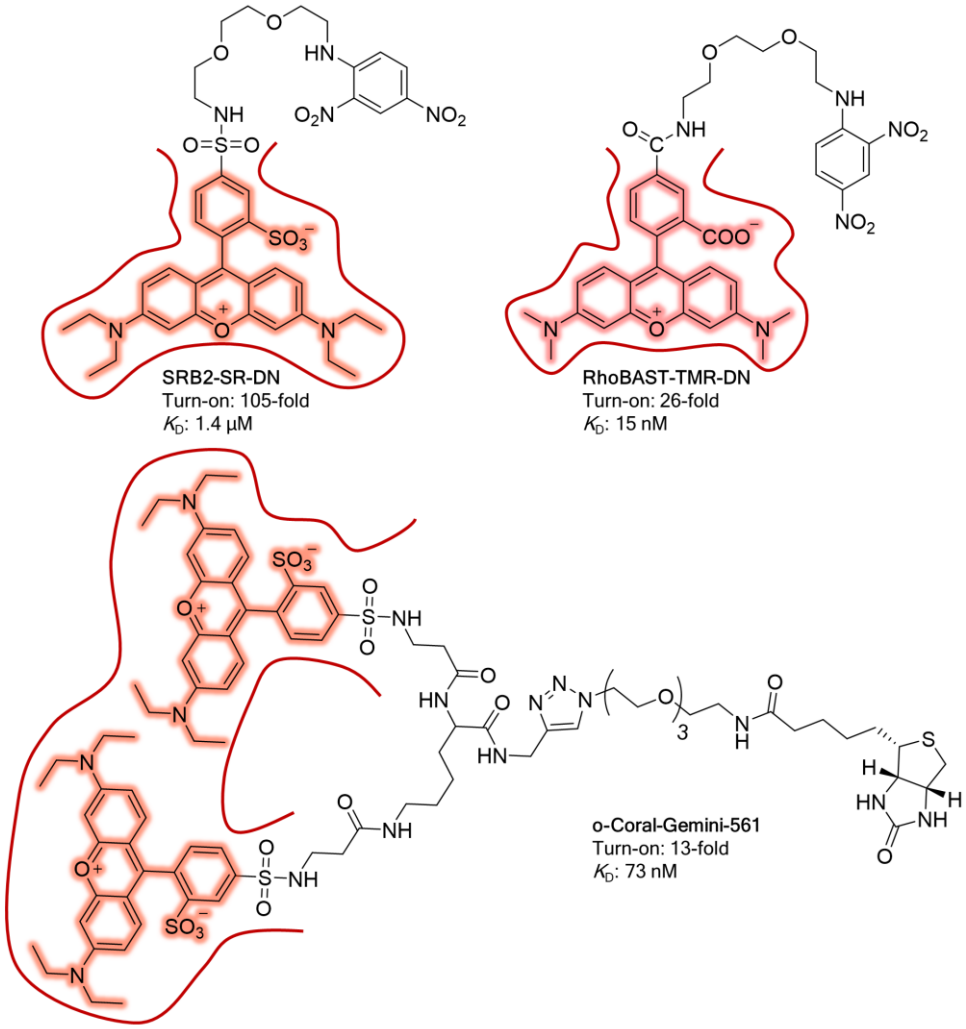


Figure 11. Rhodamine-binding aptamers (red curve) SRB2, RhoBAST, o-Coral with their bound ligands, SR-DN, TMR-DN, Gemini-561, respectively.

1.2.3.6 Quencher-binding aptamers

Instead of fluorophore-binding aptamers, quencher-binding aptamers provide more flexible and versatile ways for leveraging conventional synthetic fluorophores for RNA imaging. Synthetic fluorophores have been utilized in cellular experiments for many years. They are compatible with and have shown

good brightness and photostability in the commercially available microscopes equipped with the commonly used lasers and emission filters.
The first FLAP design based on this strategy was reported in 2011 by Uesugi and colleagues[61]. A black hole quencher (BHQ1)-binding aptamer was isolated through SELEX. BHQ1 is a universal quencher and was paired with various fluorophores (including Alexa594, Cy3, and fluorescein) and effectively quenched their fluorescence. The authors demonstrated that binding of BHQ1 to its cognate aptamer relieved the FRET-based quenching, thereby restoring the fluorescence of the fluorophore. Real-time detection of RNA transcripts in cell-free systems was achieved with this system. Further optimization by removing the sulfo group from Cy3 increased the probe's cell permeability[62]. Meanwhile, they designed an RNA-targeting (RT) aptamer-based on BHQ1-binding aptamer that only became fluorescent when hybridized to the target RNA. The probe BHQ1-Cy3 combined with the RT-aptamer enabled the specific labeling of endogenous mRNAs, including β-actin, arfaptin-2, and cortactin mRNAs, but still limited to highly abundant RNAs.
Analogously, our group reported a 78-nt dinitroaniline-binding aptamer (DNB) for RNA imaging in 2015[63]. DN as a contact quencher was conjugated to various fluorophores, generating various non-fluorescent intramolecular complexes. These fluorogenic probes all showed increased fluorescence upon DNB binding with emission ranging from green (fluorescein-DN) to red (TexasRed-DN). The best aptamer-fluorophore pair, DNB-TMR-DN (turn-on = 73-fold, K_D = 350 nM), labeled the target RNAs orthogonally with SRB2-SR-MN when co-expressed in living bacteria.
Recently, a riboswitch-based platform termed Riboglow for RNA imaging was reported (Figure 12)[64]. They utilized a natural bacterial riboswitch (102-nt) as the RNA tag and its binding small-molecule, cobalamin (Cbl), as the ligand. Cbl was appended to different fluorophores, including FAM, ATTO series, and Cy5, and functioned as an efficient fluorescence quencher. Among the multicolor substrates, Cbl-5xPEG-ATTO 590 showed the best performance, displaying a 5-fold fluorescence induction and a K_D of 34 nM upon aptamer binding. The authors demonstrated the ability of Riboglow to track β-actin mRNA as well as small non-coding RNA U1, which are not possible with the MS2 system due to its large size in living mammalian cells. However, these probes' poor cell permeability necessitates an invasive probe loading approach, thereby causing potential perturbations for live-cell imaging.

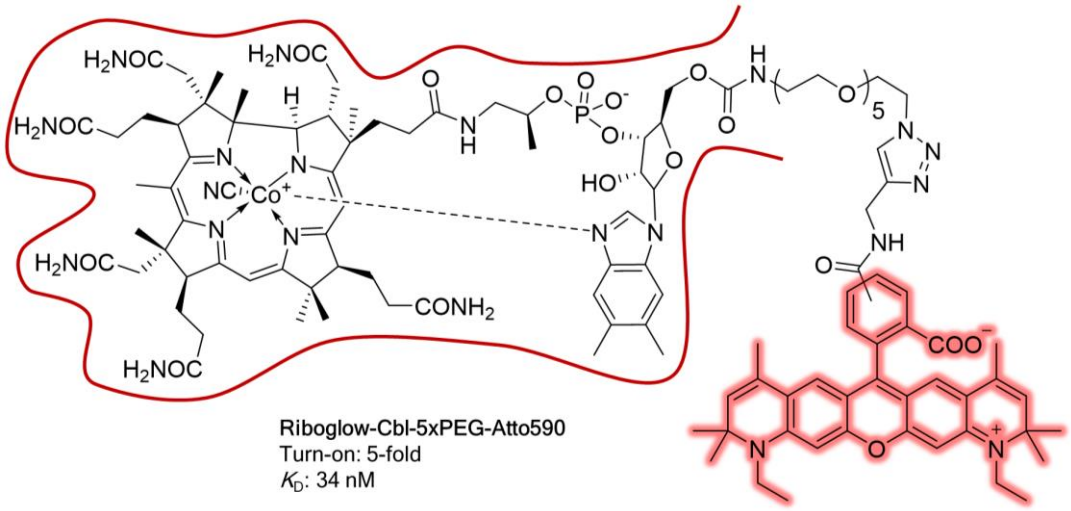


Figure 12. Cobalamin (Cbl)-binding aptamer Riboglow (red curve) with its bound ligand, Cbl-5xPEG-Atto590.

1.2.3.7 Silicon rhodamine-binding aptamer

Silicon rhodamine (SiR) is a rhodamine derivative that fluoresces in the NIR region. A 3′-carboxyl SiR was first reported for protein labeling. It became fluorescent after being covalently labeled to proteins due to the equilibrium shift of SiR from a non-fluorescent spirolactone form to a fluorescent zwitterionic form. This concept was brought to RNA labeling by developing a SiR-based FLAP system in 2019 by our group (Figure 13)[65]. A 50-nt silicon rhodamine-binding aptamer (SiRA) that selectively binds to the fluorescent ”open” form of SiR was evolved. Aptamer-bound SiR is super bright with a quantum yield of 0.98 and much more photostable than Atto 647N, a commonly used photostable dye for super-resolution stimulated emission depletion (STED) microscopy. The authors showed the first STED images of GFP mRNA using FLAPs in bacteria. Although the binding affinity (K_D = 430 nM) still has room for optimization, and this system is not yet applied to living mammalian cells, it opens up the opportunity for NIR-fluorescent super-resolution RNA imaging with high photostability and brightness.

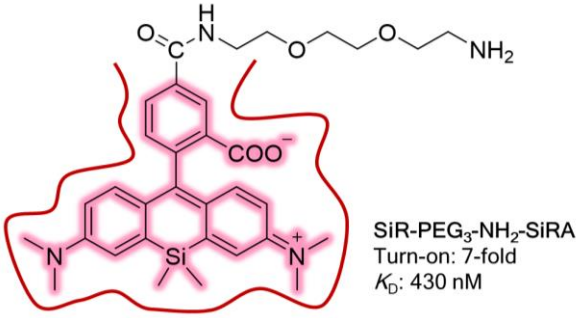


Figure 13. Silicon rhodamine-binding aptamer SiRA (red curve) with its bound ligand, SiR-PEG$_3$-NH$_2$.

Here we summarize the photophysical properties of the reported FLAPs. Some promiscuous aptamers, which bind to a palette of fluorophores, are listed with the representative ligands.

FLAP	λ_{ex} (nm)	λ_{em} (nm)	ε ($M^{-1}cm^{-1}$)	Φ	Turn-on	K_D (nM)
MGA-malachite green[39]	630[66]	655[66]	/	0.19	2360	117
Spinach-DFHBI[41]	469	501	24271	0.72	2000	537
Spinach2-DFHBI[42]	454	498	26100	0.70	770	430
Broccoli-DFHBI-1T[43]	472	507	29600	0.94	800	360
Broccoli-BI[44]	470	505	33600	0.67	1100	51
Corn-DFHO[45]	505	545	29000	0.25	417	70
Chili–DMHBO$^+$[46]	456	592	22000	0.10	/	12
Mango-TO1-Biotin[47]	510	535	77500	0.14	1100	3.9
Mango-TO3-Biotin[47]	637	658	9300	/	/	6-8
Mango II-TO1-Biotin[48]	510	535	/	/	1650	0.7 ± 0.3
Mango III-TO1-Biotin[48]	510	535	43000	0.55[67]	4400	5.6 ± 0.3
Mango IV-TO1-Biotin[48]	510	535	32000	0.41[68]	3300	11 ± 1
Pepper530[50]	485	530	65300	0.66	3595	3.5
Pepper620[50]	577	620	100000	0.58	12600	6.1
SRB2-SR-DN[56]	579	596	85200	0.65	105	1340
RhoBAST-TMR-DN[59]	564	590	96000	0.57	26	15 ± 1
o-Coral-Gemini-561[60]	580	596	14100	0.58	13	73
DNB-TMR-DN[63]	555	582	47150	0.90	73	350
DNB-SR-DN[63]	572	591	50250	0.98	56	800
Riboglow-Cbl-5xPEG-Atto590[64]	594	622	120000	0.31	5	34
Riboglow-Cbl-Cy5[64]	646	662	271000	0.26	2.7	/
SiRA:SiR-PEG_3-NH_2[65]	649	662	86000	0.98	7	430

Table 1. Characteristics of the RNA-fluorophore complexes. λ_{ex}: excitation maximum; λ_{em}: emission maximum; ε: molar extinction coefficient at maximum absorbance; Φ: quantum yield; Turn-on: fluorescence increase fold relative to the free fluorophore.

1.3 Small-molecule fluorophores for fluorescent light-up RNA aptamers

The FLAP systems use small synthetic molecules as fluorogenic ligands. Compared to fluorescent proteins, small-molecule probes have several advantages and are preferred in live-cell RNA imaging. First, they have a low molecular weight (< 900 daltons) and are therefore smaller in size, generally on the order of 1 nm[69]. Upon binding to cognate aptamers, they have less add-on weight to ROIs, which means fewer perturbations. The small size also allows

more precise visualization of the locations of ROIs, which becomes more and more important with the advances of single-molecule imaging microscopy. Second, synthetic fluorophores can possess better photophysical properties than fluorescent proteins, including higher brightness and photostability[70]. The difference can be as large as one order of magnitude. Furthermore, the properties of small-molecule fluorophores can even be optimized by rational chemical structure design.

1.3.1 Requirements of small-molecule fluorophores as FLAP ligands for live-cell imaging

Due to the superior characteristics of small-molecule fluorophores, FLAPs hold the potential for RNA imaging in living cells. When developing FLAPs, choosing the appropriate fluorophore ligands is the key to success. Several features of small-molecule fluorophores should be considered and balanced, including membrane permeability, water-solubility, absorption and emission wavelengths, brightness and photostability, and fluorogenicity.

1.3.1.1 Membrane permeability and solubility

Although numerous synthetic fluorophores have been developed, only a set of them can be used for intracellular imaging due to the poor membrane permeability and solubility (Figure 14)[69]. This is also the bottleneck problem of RNA labeling with fluorescent probes. For example, fluorophore-labeled oligonucleotides (e.g., molecular beacon) cannot pass the cell membrane unless invasive techniques are used. For FLAPs, cell-permeable fluorogenic small-molecule ligands are required for actively functioning in living cells. Dimethylindole red (DIR) and oxazole thiazole blue (OTB) are fluorogenic cyanine dyes that bind to an RNA aptamer with low-micromolar K_D values[71]. The RNA-fluorophore complexes showed good fluorescence quantum yields and excellent photostabilities *in vitro*. Nevertheless, they were limited to image a cell-surface epidermal growth factor (EGFR) receptor due to the probes' impermeability. Similarly, cobalamin conjugated fluorophores require a bead loading procedure to assist cellular uptake before performing RNA imaging[64]. Therefore, when designing or developing a new FLAP system, the ligand's membrane permeability and solubility should be carefully evaluated before starting the aptamer selection. Lipinski's Rule, which was summarized initially for the design of small-molecule orally active drugs, is also a good standard for assessing small-molecule fluorophores' biocompatibility. The ligand should follow these criteria[69]: (1) no more than five hydrogen bond donors; (2) no more than ten hydrogen bond acceptors; (3) a molecular mass lower than 500 daltons; (4) an octanol-water partition coefficient (log P) that does not exceed five. Among all the reported synthetic fluorophores, coumarins, BODIPYs, rhodamines were discovered as cell-permeable fluorophores empirically[69].

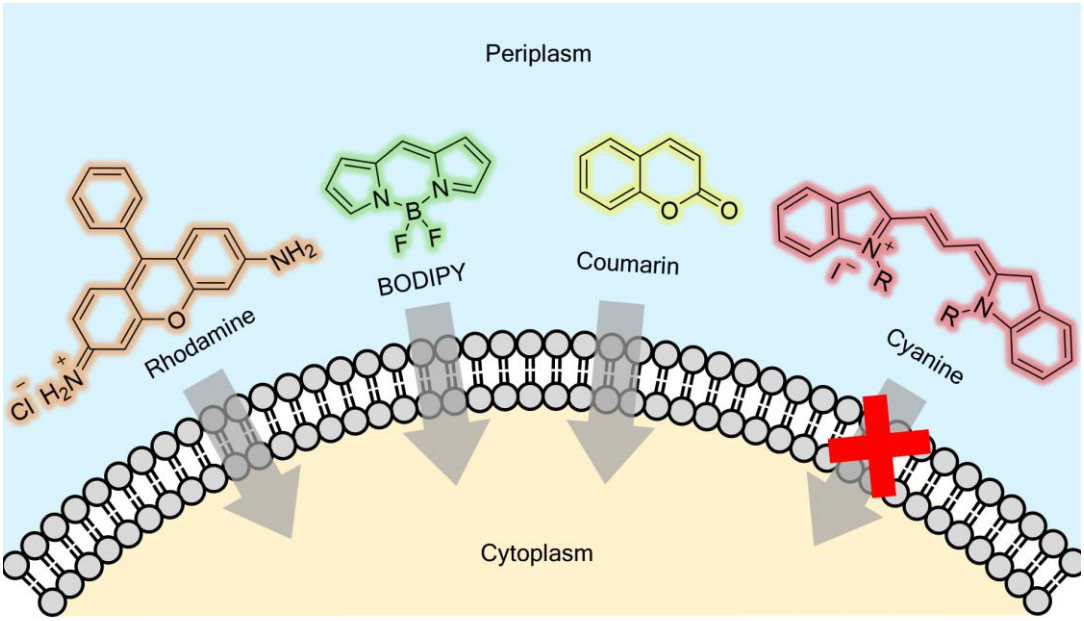


Figure 14. Passive diffusion of fluorophores into cells. The chemical structures of representative cell-permeable rhodamine, BODIPY, coumarin, and cell impermeable cyanine fluorophores are illustrated.

1.3.1.2 Absorption and emission wavelengths

Several aptamers have been reported to promiscuously bind fluorophores with different emission colors. However, the binding affinity and fluorogenicity of the aptamer to the promiscuous substrates may vary depending on the variations in the fluorophores' chemical structures. There are generally three types of promiscuous FLAPs according to the spectral-tuning methods of the cognate ligands. First, Pepper-HBC used the classic way, which tuned the spectroscopic properties of the complex by modifying the electron donor/acceptor and extending the conjugated π-system of HBC (Figure 15)[50]. The larger the conjugated π-system of the substrate is, the more red-shifted the RNA-fluorophore complex. Another emission-tuning method majorly works on the xanthene ring of rhodamine, generating rhodamine analogs (e.g., pyronin, acridine, aminoacridine, and oxazine for SRB2 aptamer[58]) or rhodamine derivatives (e.g., carbopyronin, germanium-rhodamine, and silicon-rhodamine for SiRA aptamer[65]). Unlike the methods mentioned above, quencher-binding aptamers are inborn rainbow aptamers. The quencher-appended fluorophores are flexible to change without the concern of significant affinity loss provided the existence of a quencher binding aptamer. DNB[63] and Riboglow[64] are good examples based on this strategy. Notably, when it comes to live-cell or *in vivo* RNA imaging, absorption and emission wavelengths in the NIR window are particularly favorable due to reduced autofluorescence and decreased phototoxicity of cells at this wavelength range[72-74].

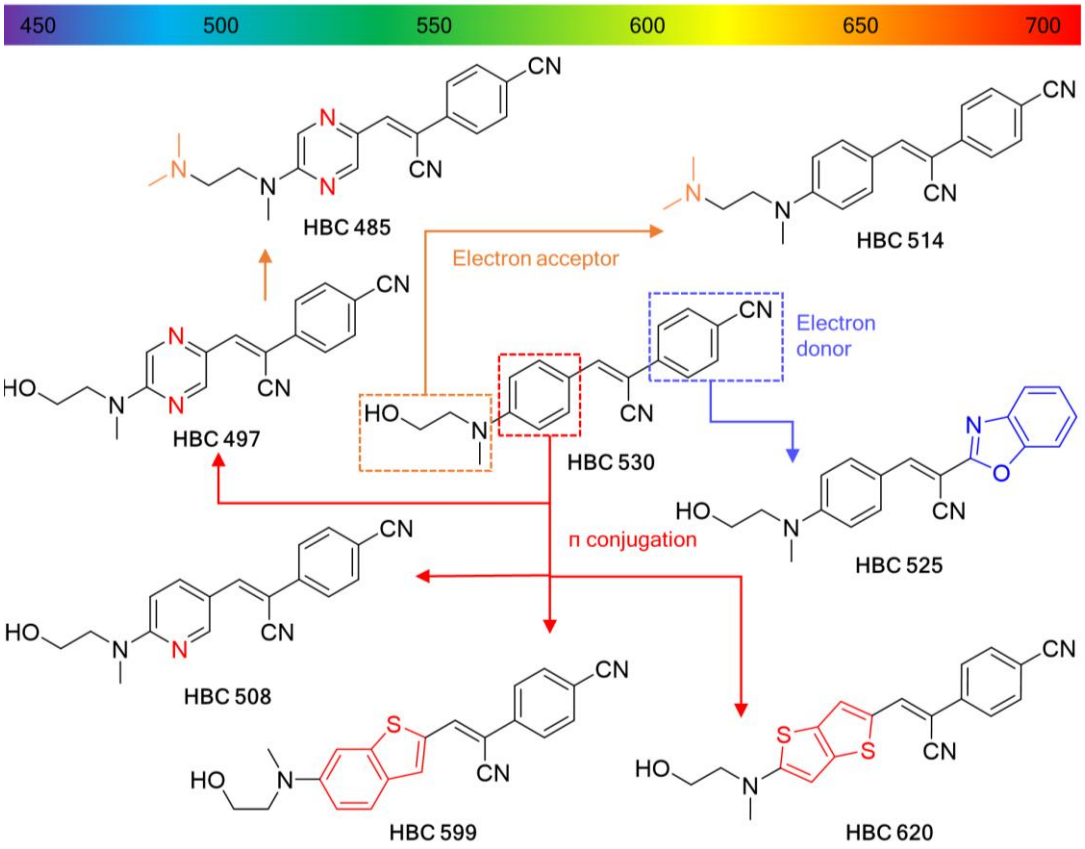


Figure 15. Substrate promiscuity of Pepper-HBC (adapted from Chen et al. 2019[50]). The chemical structures of HBC derivatives are shown with the emission maxima indicated in their names. The modification sites are highlighted in orange, red, and blue according to their functions.

1.3.1.3 Brightness and photostability

Two important parameters for characterizing a fluorophore are the molar extinction coefficient (ε) and quantum yield (Φ). Molar extinction coefficient represents how much light is absorbed by a fluorophore at a given wavelength, while quantum yield describes how many photons are emitted from a fluorophore relative to photons absorbed. Brightness is the multiplication of the two parameters ($\varepsilon \times \Phi$), a measure of fluorescence output per molecule. Higher brightness means less dye or lower laser power is needed to obtain the same number of photons, hence lower cytotoxicity. In addition to that, the higher brightness of a FLAP allows fewer required aptamer tags for an ROI, which potentially lowers the risk of altering its function/location and provides a more precise representation of its location.

Photostability describes the fluorescence performance of a fluorophore to light irradiation. High photostability means the fluorescence signal of a fluorophore is stable under continuous illumination without photobleaching over time. Spinach-DFHBI is considered not photostable since DFHBI undergoes photoinduced isomerization and subsequently dissociates from Spinach aptamer under illumination, leading to fast fluorescence decay[75]. Pepper substrates

(HBC497, HBC508, HBC514, HBC525, and HBC530) also showed similar phenomenons with fast and reversible photoswitches[50]. In contrast, small-molecule fluorophores such as rhodamines, coumarins, cyanines, and BODIPYs demonstrated reasonable photostability in imaging various targets[69]. For example, one rhodamine derivative (TMR-DN), as the substrate for RhoBAST aptamer, exhibited excellent photostability in RNA imaging[59].

1.3.1.4 Fluorogenicity

Fluorogenicity in FLAPs refers to the ability of a fluorophore to increase fluorescence upon interacting with its cognate aptamer. Fluorogenic probes significantly reduce the background signal from the unbound dye, thereby increasing the signal-to-noise ratio. Three types of fluorogenic probes are used as ligands for FLAPs to get background-free images. The first category is the rotational/vibrational-quenched small molecules such as HBIs (ligands for Pepper)[41], TOs (ligands for Mango)[47], and HBCs (ligands for Pepper; Figure 16A)[50]. The excited fluorophores relax to the ground states with vibrational energy release rather than emission photons, leading to low fluorescence. The binding of the small-molecules to their cognate aptamers restrains the molecular motions and restores the fluorescence. The second type is fluorophore-quencher conjugates (Figure 16B). As indicated by the name, the fluorophore's absorbed energy was released via the non-radiative decay from the quencher. The direct interaction between the fluorophore/quencher and the cognate aptamer disrupt the non-fluorescent intramolecular complex and induce fluorescence. Sulforhodamine B-quencher (ligands for SRB2)[56] and Cbl-fluorophore (ligands for Riboglow)[64] are representative examples of this class. The third group is the spirocyclization-based fluorophores (Figure 16C). They exist in the non-fluorescent spirocyclic form in the unbound state. Complexing with the cognate aptamers switches the probes to the fluorescent zwitterionic form, resulting in fluorescence enhancement. SiR-PEG_3-NH_2 (ligand for SiRA) is a good representative[65].

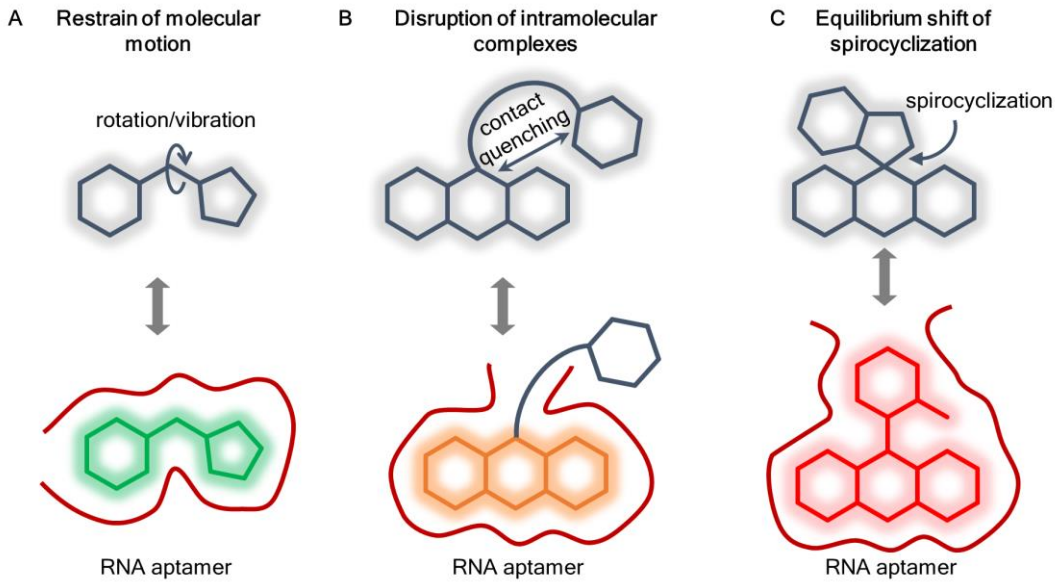


Figure 16. Fluorogenicity mechanism of FLAPs. (A) Small molecules are non-fluorescent in solution due to the rotational/vibrational motions. The binding of the small molecule to the cognate aptamer restrains the motion and restores fluorescence. (B) Fluorophore-quencher conjugates are non-fluorescent complexes. The specific interactions between the fluorophore/quencher and the cognate RNA aptamer break up the intramolecular complex, resulting in fluorescence enhancement. (C) The non-fluorescent spirocyclic probes bind to the cognate RNA aptamers, shifting the equilibrium to the fluorescent "open" form.

It is worth noting that although some FLAPs showed extraordinary fluorogenic properties *in vitro*, the live-cell imaging only shows modest contrast. Besides the improper aptamer folding in cells, one big limitation comes from the unspecific background staining of the probes inside cells. Some fluorophores are found to intercalate into proteins and other cellular components or accumulate in specific organelles[69]. These unspecific bindings induce background fluorescence, disturbing the specific signal detection. The intracellular performances of the fluorogenic ligands need to be taken into considerations before developing new FLAPs.

1.3.2 Rhodamine-based near-infrared fluorophore

NIR imaging has the advantages of deeper penetration, less autofluorescence, and low phototoxicity due to the low absorption of cells and tissues in this wavelength range[72, 74]. Even though many NIR fluorophores have been reported, most of them are not well suited for live-cell RNA imaging as FLAP ligands. For example, TO-3-Biotin (ligand for Mango)[48], malachite green (ligand for MGA)[66], and Cbl-Cy5 (ligand for Riboglow)[64] showed good fluorogenic behaviors *in vitro*. However, their performances in living cells have not been investigated yet (TO-3-Biotin) or were less satisfying (malachite green and Cbl-Cy5). This is probably due to the high cytotoxicity, low cell-permeability, and poor photophysical properties (e.g., low brightness and photostability) of these probes.
Rhodamines are a family of related dyes that all contain the pyronine structure as the chromophore. They have been tailored as fluorescent sensors, organelle

markers, and labeling agents for various chemical and biological applications. They possess good cell-permeability, photostability, and brightness, making them greatly favorable for live-cell imaging. Two representative NIR fluorophores, a rhodamine derivative silicon rhodamine and a rhodamine-inspired fluorophore benzopyrylium-coumarin, will be discussed in the following parts.

1.3.2.1 Silicon rhodamine

SiR is derived from the classic rhodamine by replacing the oxygen atom with a silicon moiety of the pyronine structure. The substitution of oxygen with silicon results in a relatively low-lying lowest unoccupied molecular orbital (LUMO) level of the chromophore, thereby resulting in a bathochromic shift of the excitation and emission wavelengths. Koide *et al*. in 2011 first reported the synthesis of a 3-methyl SiR. It demonstrates NIR fluorescence with a high quantum yield and excellent photostability, comparable to those of the original rhodamine[76]. In 2012, Wang *et al*. reported the synthesis of a 3′-carboxyl SiR and its amide derivatives[77]. The amide coupling turned 3′-carboxyl SiR into completely "dark" spirolactams. The ring-opening of spirocyclic SiR after reacting with the target analyte (i.e., Hg^{2+}) gave rise to a strong NIR emission. Nevertheless, the sensing is, unfortunately, irreversible.
The importance of the reversible spirolactonization of 3′-carboxyl SiR became apparent with the introduction of 3′-carboxyl SiR for protein[73], cytoskeleton[78], and DNA[79] labeling by Lukinavicius *et al*. (Figure 17). One unique feature of the spirocyclization-based rhodamines is that they possess a dynamic equilibrium between a non-fluorescent spirolactone and a fluorescent zwitterion. The silicon substitution of rhodamine reduced the electron density of the xanthene. Therefore, its propensity to form the non-fluorescent spirolactone in aqueous solutions was largely increased after the introduction of a carboxyl group to the 3′-position of the upper benzene ring of SiR. The spirolactone can more readily across the cell membrane compared to its zwitterionic counterpart, that is to say, higher cell permeability of the carboxylated SiR. The binding of 3′-carboxyl SiR-based probes to their respective targets shifted the equilibrium from the non-fluorescent spirolactone to the fluorescent zwitterion, generating fluorescence enhancements. Unlike conventional fluorophores, which generally show unspecific staining of cellular structures due to the probes' hydrophobicity, the unspecific binding of the spirocyclic SiR primarily keeps non-fluorescent, showing a reduced background signal in living cells. The bright signal upon binding with targets combined with the reduced background signal enables 3′-carboxyl SiR to be a popular fluorogenic probe for high-contrast fluorescence imaging in living cells.

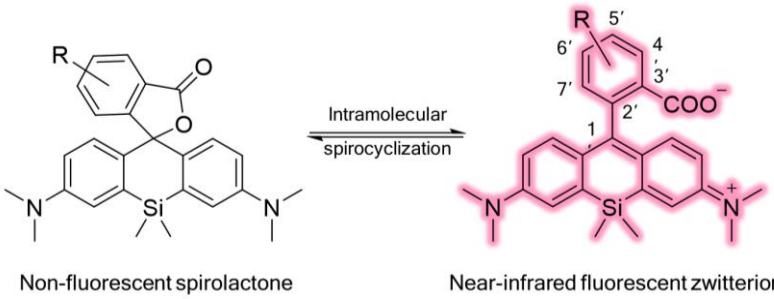


Figure 17. Environment-dependent intramolecular spirocyclization of silicon rhodamines. R depicts the position where the fluorophores are substituted for labeling purposes.

As mentioned above, our group developed a SiR-binding aptamer SiRA for bacterial mRNA imaging with excellent brightness and photostability. However, SiR only showed a moderate (7-fold) fluorescence turn-on upon aptamer binding, limited mainly by the relatively low spirolactone/zwitterion ratio of SiR in aqueous solutions. Recently, new SiR derivatives such as HMSiR[80], JF635[81], and MaP700[82] have shown higher degrees of spirolactonization with improved fluorogenicity abilities, which might be better candidates for generating the next-generation of SiR-based NIR FLAPs.

1.3.2.2 Benzopyrylium-coumarin

Benzopyrylium-coumarin (BC) is a set of hybrid fluorophores of benzopyrylium and coumarin. The first BC fluorophore, 2-(7-Diethylamino-2-oxo-2H-1-benzopyran-3-yl)-4-(2-carboxyphenyl)-7-diethylamino-1-benzopyrylium, was reported in 1995 as a laser dye[83]. Like the carboxyl-rhodamine, the carboxyl-BC possesses a reversible intramolecular spirocyclization that makes it exits in a dynamic equilibrium between a spirocyclic "closed" form and a zwitterionic "open" form. Unlike the carboxyl-rhodamine, which shows an "on-off" fluorescence switch, the carboxyl-BC displays a "blue-NIR" fluorescence switch between the two configurations. The spirocyclic form emits blue fluorescence from the coumarin structure, while the zwitterionic form exhibits NIR fluorescence from the extended π-conjugated system. Liu *et al*. employed the dual-emission nature of BC fluorophores and developed a ratiometric Hg^{2+} sensor based on this platform (Figure 18)[84]. The free probe displayed an absorption and an emission that are typical of that of a 7-diethylaminocoumarin moiety. However, with the addition of Hg^{2+}, the absorption and emission intensity at the low wavelength range reduced while new absorption and emission bands appeared at the NIR window. *In vitro* study further demonstrated the ratiometric responses (F_{695}/F_{480}) of the BC fluorophore to Hg^{2+} concentrations (0-30 mM) in an excellent linear manner. Compared to single-color fluorogenicity, ratiometric sensing employs the fluorescence intensity ratio of two emissions. Variability arising from the uneven cellular uptake of the dye, heterogeneous distribution of the probe inside cells, different environmental conditions (e.g., pH, polarity, and temperature) are therefore canceled out by

using the ratiometric signal[85-86]. Later, a series of BC-based sensors for ratiometric sensing of different analytes, including cysteine and homocysteine[87], hypochlorite[88], hydrogen sulfide[89], hydrazine[90], hydrogen peroxide[91], and singlet oxygen[92] were developed and used in various biological contexts.

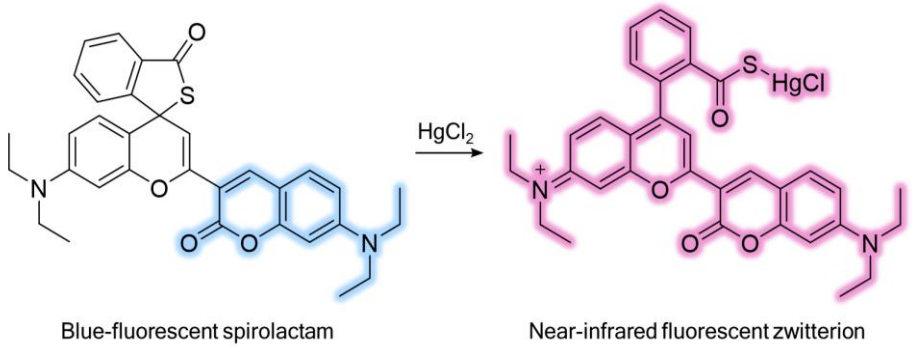


Figure 18. Benzopyrylium-coumarin fluorophore as a ratiometric probe for Hg^{2+} sensing. The BC thiolactam reacts with Hg^{2+}, opening up to a NIR-fluorescent zwitterion.

Due to the unique dual-color feature, BC fluorophores are interesting candidates for developing a color-shifting NIR FLAP for quantitative RNA imaging and monitoring. However, unlike the 3′-carboxyl SiR, which favors the non-fluorescent spirocyclic form and gives rise to NIR fluorescence upon aptamer binding, the 3′-carboxyl BC predominantly stays in the NIR-fluorescent zwitterionic form in aqueous solutions. In order to get a NIR fluorescence turn-on upon aptamer binding, BC derivatives that stay in the blue-fluorescent form, however, can be specifically switched by the aptamer to the NIR-fluorescent form are needed.

1.3.3 Fine-tuning of the spirocyclization-based fluorophore

Rhodamine fluorophores that utilize the intramolecular spirocyclization-based fluorogenicity have shown excellent live-cell imaging performances due to the less unspecific staining and higher cell permeability[73, 78-79]. However, most reported rhodamines present in the zwitterionic fluorescent form in aqueous solutions, making them impractical to label biomolecules as fluorogenic probes directly. Fine-tuning of the intramolecular spirocyclization equilibrium is therefore needed. In the case of developing spirocyclization-based NIR FLAPs, the key challenge is to turn those non- or less fluorogenic NIR fluorophores into highly fluorogenic ligands. The ideal scenario is that by controlling the intramolecular equilibrium, the fluorophore predominantly stays in the non-NIR-fluorescent (can be non-fluorescent or fluoresce in other colors) spirocyclic form. Only when binding to the target aptamer, the NIR fluorescence generates from the zwitterionic form due to the equilibrium shift.
Currently, two principles have been developed to obtain spirocyclization-based fluorogenic probes using 3′-carboxyl rhodamine as the core structure (Figure 19)[80]. Increasing either the electrophilicity of the chromophore (i.e., the xanthene

moiety) or the nucleophilicity of the intramolecular nucleophile (i.e., the carboxyl group) facilitates a more efficient intramolecular nucleophilic attack of the nucleophile to the spiro-carbon (C9) of the xanthene. Therefore, an equilibrium shift of the rhodamine fluorophore to the spirocyclic form is achieved. However, the parameters should be carefully tuned that the fluorophore is not completely converted to the spirocyclic form. The target-binding should be able to reversibly open up the spirolactone.

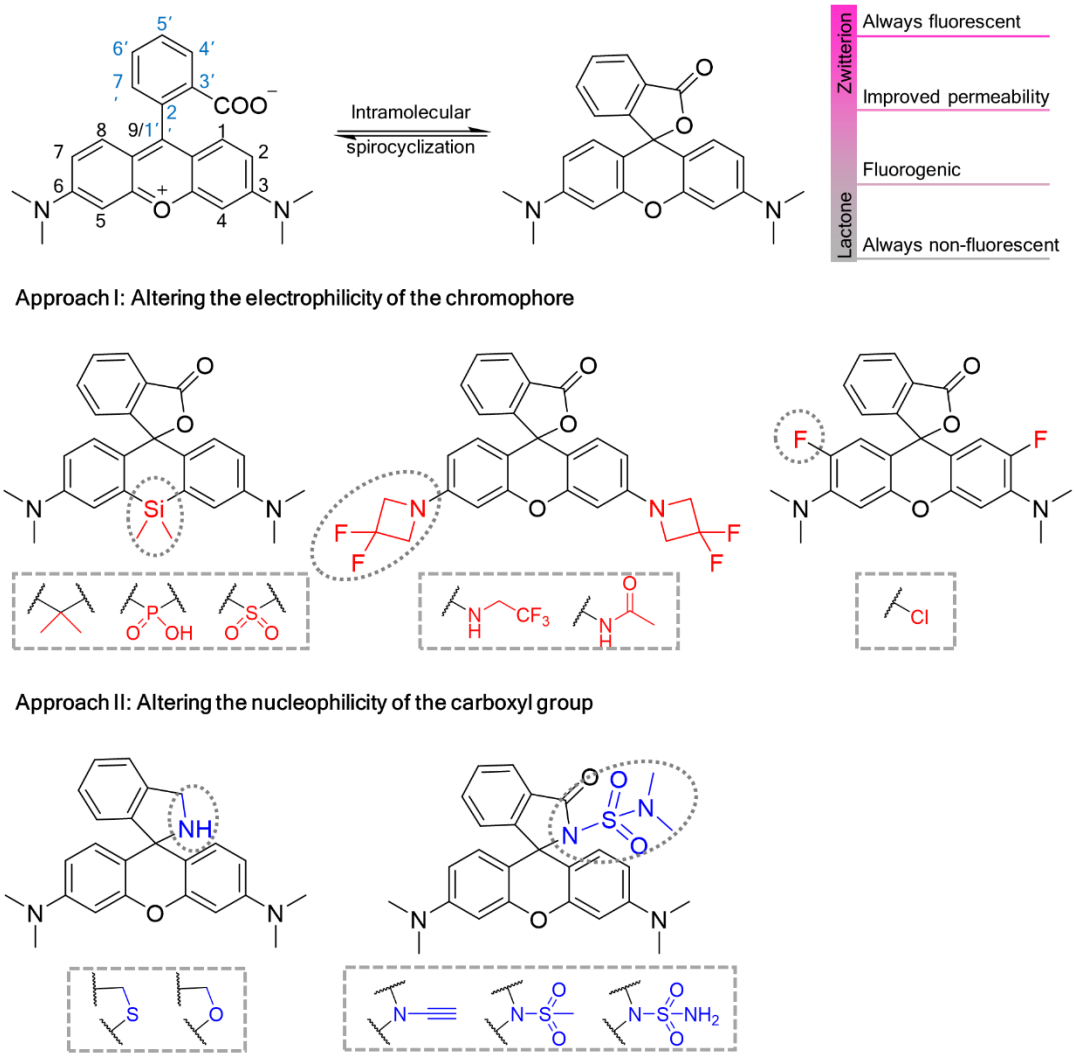


Figure 19. Fine-tuning of the spirocyclization-based rhodamine fluorophore. Increasing either the electrophilicity of the chromophore or the nucleophilicity of the intramolecular nucleophile shifts the equilibrium to the spirocyclic form. The modification sites are indicated with dashed circles/boxes.

1.3.3.1 Increasing the electrophilicity of the chromophore

The first approach to tune the electron density of the rhodamine chromophore is to replace the central oxygen atom of the xanthene moiety with group 14 (e.g., carbon, silicon, and germanium,), group 15 (e.g., phosphorous), or group 16 (e.g., sulfur, selenium, and tellurium) elements[93-94]. The resulted rhodamine

derivatives not only show increased spirocyclic form in water but also shifted emissions in deep-red and NIR regions that are favored in bioimaging. The most famous example derived from this design strategy is the silicon-substituted rhodamine, 3′-carboxyl SiR, a fluorogenic probe used for various biomolecule targets.

Another common way to modulate the electrophilicity of the chromophore is by introducing electron-withdrawing groups to the N,N-dimethylamino group of rhodamine, which is proven to efficiently shift the equilibrium of the fluorophore to the "closed" lactone form in addition to a hypsochromic shift in both absorption and emission. The difluoroazetidinyl-substituted rhodamines (oxygen-, carbon- and silicon-) all showed increased fluorogenicity compared to their respective parent structures upon protein binding[81]. Introduction of the 2,2,2-trifluoroethyl group at the nitrogen atom of rhodamines also favored the spirocyclic form. The corresponding derivatives enabled the specific labeling of cytoskeletal filaments with high contrast in live-cell imaging[95]. Besides, the acetylated-amino group was reported to stabilize the spirocyclic structure. The Urano group developed a series of acetylated-rhodamines that stay in the non-fluorescent spirocyclic form in aqueous solutions and become fluorescent after sensing specific enzymes (e.g., leucine aminopeptidase, fibroblast activation protein, and β-galactosidase)[96-97].

The last method is to destabilize the positive charge of the xanthene ring. Halogenation of the xanthene fragment at 2- and 7-positions promoted the spirocyclic form, which exhibits fluorogenic responses towards biomolecule labeling and sensing[95, 98].

1.3.3.2 Increasing the nucleophilicity of the carboxyl group

The zwitterionic 3′-carboxyl rhodamine uses the carboxyl group as the nucleophile for the intramolecular spirocyclization reaction. When stronger nucleophiles such as hydroxymethyl, aminomethyl, and mercaptomethyl were used, they tended to shift the equilibrium of rhodamines to the spirocyclic form by different extents according to their varied nucleophilicities (mercaptomethyl > aminomethyl > hydroxymethyl)[80]. Among all the derivatives, the hydroxymethyl silicon rhodamine (HMSiR) was discovered as a spontaneously blinking fluorophore (fluorescent zwitterionic form constitutes 1% of total amount) and used for live-cell SMLM[80].

Analogously, by chemical alteration of the carboxyl group to amides, rhodamine derivatives have been widely used as fluorogenic probes for sensing different analytes, including pH[99], metal ions[100], and reactive oxygen species (ROS)[101]. In 2020, Wang *et al*. reported a strategy using electron-deficient amides for developing reversible fluorogenic rhodamine probes. These probes are non-fluorescent in the spirocyclic form while binding to the biomolecules activates their fluorescence due to the equilibrium shift towards the zwitterionic form. Using these probes, they demonstrated wash-free, multicolor, live-cell labeling of proteins and other cellular targets using super-resolution microscopy[82].

We only discussed how to tune the zwitterionic fluorescent rhodamines into the non-fluorescent spirocyclic forms in the above texts. There are cases that they become permanently non-fluorescent after chemical modification. Some methods are reported based on tuning the upper phenyl ring to rescue these probes for biological use. For example, Grimm *et al*. successfully tweaked the non-fluorescent sulfone-substituted rhodamine into a fluorogenic probe by halogenating the upper phenyl ring, presumably by lowering the pKa of the benzoic acid moiety[93].

It is worth noting that a single approach alone sometimes is not sufficient to tune the spirocyclization equilibrium to a satisfying extent. According to different experimental purposes, a combination of two or more approaches is applied to modulate the overall spirolactone/zwitterion ratio of the probe. By fine-tuning the spirocyclization-based fluorophores, a large variety of fluorogenic probes can be developed and used as ligands for developing FLAPs with high specificity and fluorogenicity.

2 Motivation

RNA plays an important role in gene expression and regulation[2]. Visualizing the intricate RNA dynamics inside living cells would help us understand the mechanisms behind cell development, differentiation, and programmed death[102]. Among all the RNA imaging techniques, FLAPs represent emerging and exciting possibilities for live-cell RNA tracking due to the small tag size, superior fluorescence properties, and ease of operation. An aptamer tag is a short RNA sequence that specifically binds to a cell-permeable small-molecule probe and activates its fluorescence. The aptamer tag can be genetically fused to an RNA of interest in a tandem array fashion. Thus, the RNA of interest can be visualized by resolving the fluorescent aptamer-fluorophore complexes using microscopy.
Imaging in the NIR window is preferred for biological systems due to the low photocytotoxicity and autofluorescence of cells and tissues in this wavelength range[72, 74]. Although a set of FLAPs have been developed, FLAPs that fluoresce in the NIR range for live-cell imaging are still rare[65, 71]. Therefore, in this thesis, we aim to develop two new NIR FLAPs with high specificity and sensitivity for RNA imaging in living cells. Among all the reported fluorophores, spirocyclization-based fluorogenic probes have shown great advantages in background-free labeling of various cellular targets in living cells[81-82]. They exhibit high cell permeability and low background staining inside cells. However, this spirocyclization-based fluorogenicity has not been rationally exploited yet in the RNA labeling field. Therefore, two NIR fluorophores, benzopyrylium-coumarin and silicon rhodamine, both possessing the intramolecular spirocyclization equilibrium, were selected as the NIR fluorogenic ligands to develop new NIR FLAPs.

2.1 Benzopyrylium-coumarin binding color-shifting aptamer-fluorophore module

Currently, all the reported FLAPs, to our knowledge, are based on single-wavelength fluorescence readout. The different cellular uptake of the probe, the morphology of the cells (e.g., cell thickness), the environmental conditions during the imaging (e.g., temperature) would cause signal fluctuation, thus introducing inaccuracy during imaging analysis[85-86]. Ratiometric imaging as a self-calibration system is preferred in many biological applications since the simultaneous detection of two emissions can cancel out the varieties.
The first goal of this thesis is to develop a color-shifting aptamer-fluorophore module for ratiometric RNA imaging. The BC fluorophore is a rhodamine-inspired environmentally sensitive fluorophore, which exists in an intramolecular equilibrium between a cyan-fluorescent spirocyclic form and a NIR-fluorescent zwitterionic form[84]. The BC fluorophore can be adjusted to primarily emit cyan fluorescence in the spirocyclic form, whereas complexing with the cognate RNA aptamer induces the formation of NIR-fluorescent zwitterionic form (Figure 20). The two emissions from either unbound BC or

aptamer-BC complex could be used for imaging and evaluating the abundance of aptamer-tagged RNAs of interest.

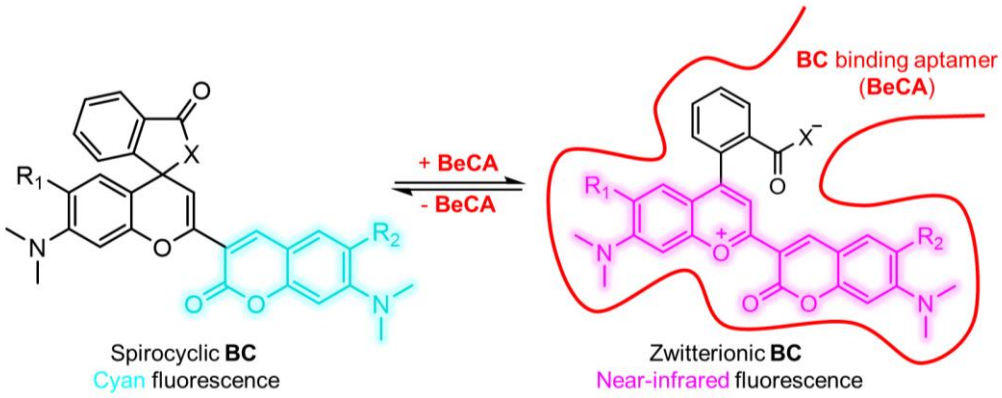


*Figure 20. BC-binding color-shifting NIR FLAP for RNA imaging. Unbound BC is cyan-fluorescent, whereas **BeCA**-bound BC is NIR-fluorescent.*

The workflow of developing such a color-shifting NIR FLAP is shown (Figure 21). First, a zwitterionic BC fluorophore would be synthesized and used as a bait for *in vitro* RNA aptamer selection (SELEX). After activity screening of the isolated RNA sequences, high-affinity and bright binders would be identified. The truncation and mutation study of the best aptamer candidate would be performed to get a minimal BC zwitterion-binding aptamer. Next, rational chemical modifications would be applied to generate BC analogs with diverse zwitterion/spiro-lactone(lactam) ratios. Screening of these color-shifting ligand candidates would offer the best aptamer-fluorophore pair with a large dynamic range and a high binding affinity. Finally, the optimized color-shifting aptamer-fluorophore module would be tested for the proof-of-principal ratiometric RNA imaging in living cells.

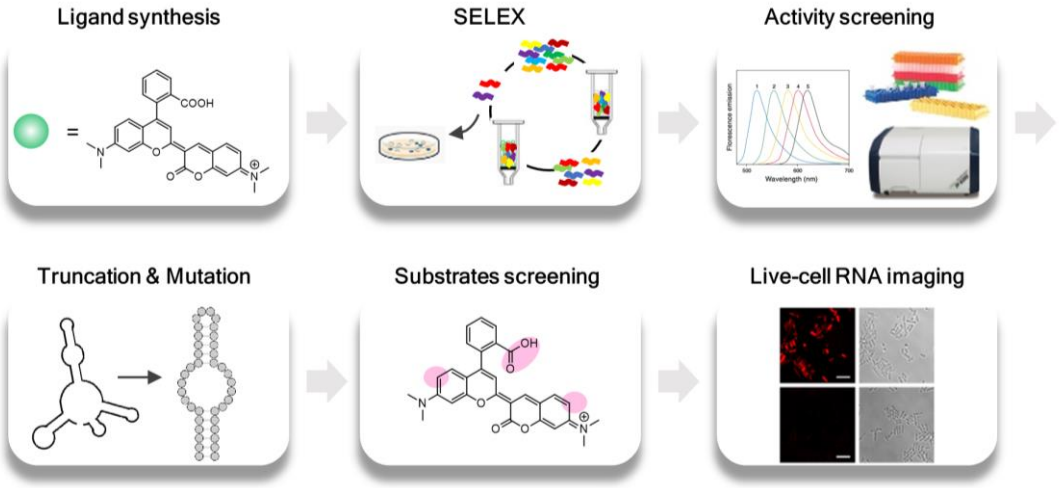


Figure 21. The workflow of developing a color-shifting NIR FLAP using BC fluorophores as ligands.

2.2 Silicon rhodamine binding fluorescent light-up aptamer

Spirocyclization-based fluorogenic probes (e.g., SiR, JF, and MaP dyes) are interesting ligands for developing new FLAPs due to their excellent cell permeability and photophysical properties[73, 81-82]. However, traditional affinity-oriented aptamer selection methods (SELEX and its variants) are not ideal for generating FLAPs for this kind of fluorophores. Enrichment of non-fluorescent spiro-lactone(lactam)-binding sequences instead of fluorescent zwitterion-binding sequences would be obtained since these fluorophores as selection baits exist primarily in the spirocyclic form under *in vitro* selection conditions. In contrast, fluorescence-activated cell sorting (FACS) is a good alternative aptamer selection method as both affinity and brightness of a FLAP can be assessed during selection. The workflow of developing a FACS-based platform for evolving new spirocyclization-based FLAPs is shown (Figure 22). A standard protocol for sorting fluorescent bacteria would be established first. Then, the bacteria library expressing various aptamers would be sorted according to their fluorescence signal in the presence of a fluorogenic ligand. The highly fluorescent cells would be collected and used for the next round of sorting. After several rounds with increased selection pressure, the enriched RNA sequences from the highly bright bacterial cells would be sequenced and analyzed. *In vitro* characterization and truncation would reveal the minimal RNA motif (aptamer) that specifically binds and activates the fluorescence of the small-molecule ligand.

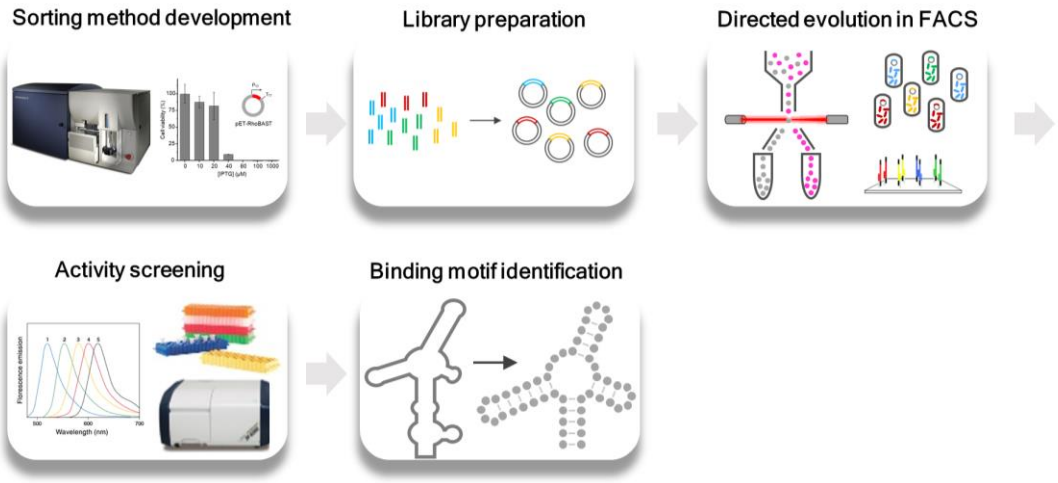


Figure 22. The workflow of developing a spirocyclization-based FLAP through FACS.

Recently, a silicon rhodamine-binding aptamer (SiRA) was obtained via SELEX and has been successfully utilized for live-cell RNA imaging[65]. Nevertheless, mammalian cell imaging of RNAs of interest using SiRA is not yet possible, presumably due to the poor binding affinity between SiR and SiRA (K_D = 430 nM).

Therefore, the second goal of this thesis is to develop a FACS-based aptamer selection platform and use this platform to evolve a new SiR-binding NIR FLAP

with improved detection sensitivity. The newly evolved aptamer should bind to SiR with an improved binding affinity (ideally down to a low-nanomolar range) while retaining the outstanding brightness and photostability of SiRA (Figure 23). The new aptamer-SiR module holds the potential to label RNAs of interest in mammalian cells.

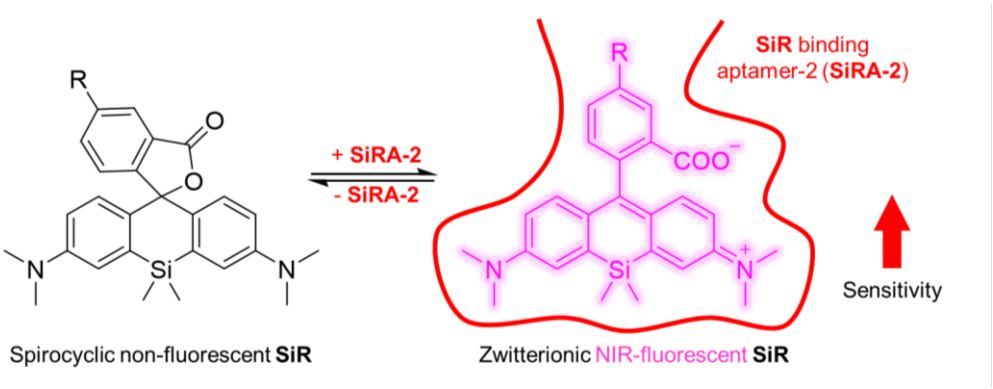


*Figure 23. An improved SiR-binding NIR FLAP. Unbound spirocyclic SiR is non-fluorescent while binding to the SiR-binding aptamer-2 (**SiRA-2**) shifts SiR to the zwitterionic form and turns on the NIR fluorescence.*

3 Results and discussion

In this thesis, two NIR FLAPs for RNA imaging were developed (chapter 3.1 and chapter 3.2, respectively). The majority of the content described in chapter 3.1 was covered in an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al.* (in preparation). The text was originally written by myself.

3.1 Development of a benzopyrylium-coumarin binding color-shifting near-infrared fluorescent light-up aptamer

Inspired by the unique dual-color feature of BC fluorophores (chapter 1.3.2.2), a BC-binding color-shifting NIR FLAP was developed through *in vitro* aptamer selection and applied for live-cell RNA imaging. All the results will be described and discussed in detail in the following texts.

3.1.1 Synthesis and characterization of BC1-based ligands

BC fluorophores exist in an equilibrium between a cyan-fluorescent spirocyclic form and a NIR-fluorescent zwitterionic form. To evolve a color-shifting NIR FLAP module, an RNA aptamer that activates the NIR fluorescence of BC is required. In other words, the RNA aptamer should possess high specificity and binding affinity towards the NIR-fluorescent zwitterionic form rather than the cyan-fluorescent spirocyclic form of BC. **BC1** is the tetramethyl version of the reported hybrid fluorophore of benzopyrylium and coumarin [84]. It mainly exists in the NIR-fluorescent zwitterionic form in buffer solution. Therefore, **BC1** was conjugated to solid supports through a flexible polyethylene glycol (PEG) linker and used as a bait for the aptamer selection (Figure 24).

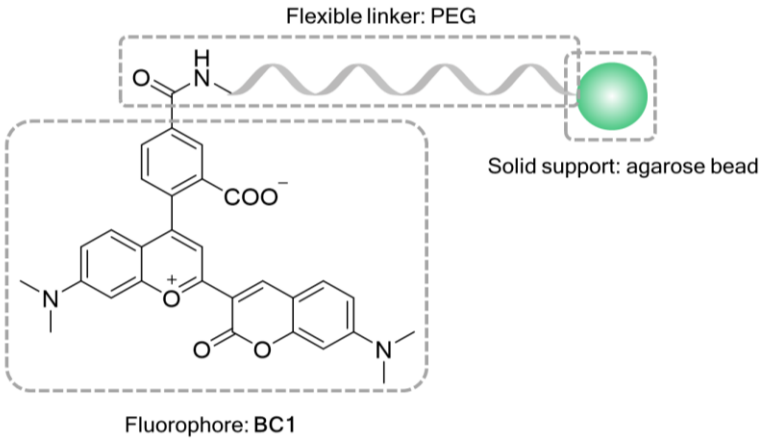
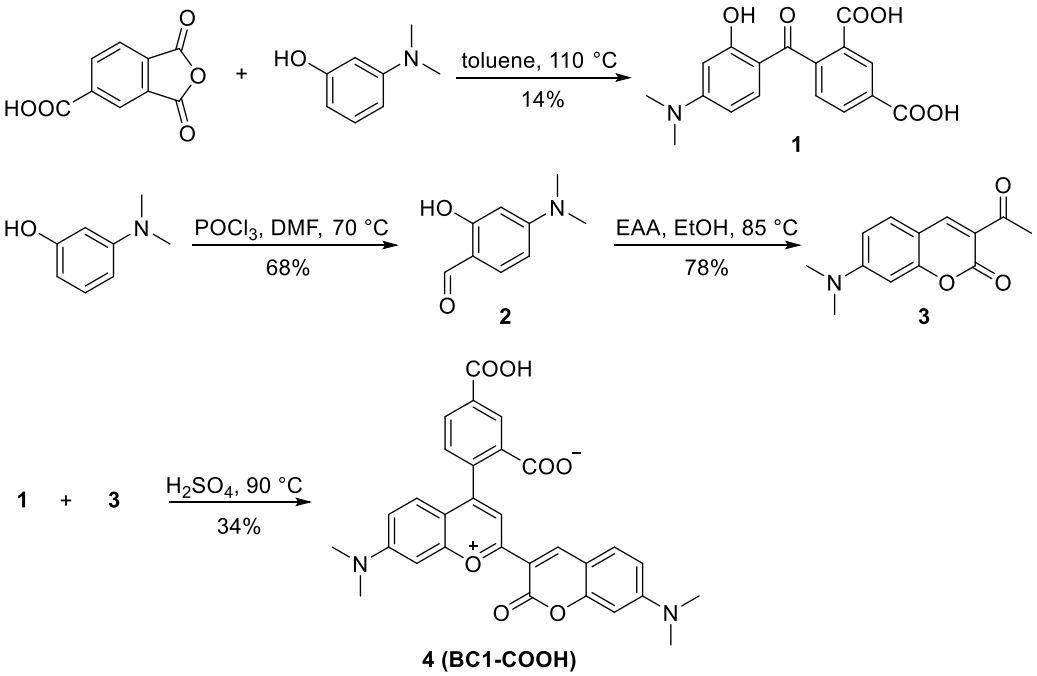


Figure 24. ***BC1*** *attached to the solid support via a PEG linker.*

3.1.1.1 Synthesis of BC1-COOH

For linker conjugation, we introduced an extra carboxylic acid group to the 5′-position of the upper benzene ring of **BC1**. The synthetic routes of the carboxylated **BC1** (**BC1-COOH**) were adapted from the reported literature (Scheme 1)[84]. Trimellitic anhydride reacted with dimethylaminophenol to give 5′-isomer (**1**), which was isolated from the 6′-isomer by crystallization in methanol with a yield of 14%. In parallel, the coumarin (**3**) was readily prepared via a *Vilsmeier-Haack* reaction of dimethylaminophenol followed by a *Knoevenagel* condensation reaction with acetoacetate in a total yield of 53%. The two intermediates **1** and **3** were condensed to the target compound **BC1-COOH** (**4**) with two carboxylic groups: the 3′-carboxyl group acts as the molecular switch for the intramolecular spirolactonization and the 5′-carboxyl group bridges the linker and the fluorophore.



*Scheme 1. Synthesis of **BC1-COOH** (**4**)*

3.1.1.2 Spectroscopic characterization of BC1-COOH

The spectroscopic properties of **BC1-COOH** were characterized. **BC1-COOH** exhibited the typical environmentally sensitive intramolecular spirocyclization in the solvent-dependent absorption measurement, as expected (Figure 25A). It remained as spirolactone in dimethylformamide (DMF; $\lambda_{spirolactone\text{-}abs}$ = 422 nm) while the equilibrium shifted to the zwitterionic form ($\lambda_{zwitterion\text{-}abs}$ = 638 nm) in methanol (MeOH; Figure 25B). Next, we examined the absorption and emission spectra of **BC1-COOH** in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (Figure 25C). **BC1-COOH** showed strong absorbance in the long-

wavelength region and emitted NIR fluorescence, demonstrating that **BC1-COOH** exists predominantly in its zwitterionic form in aqueous solutions.

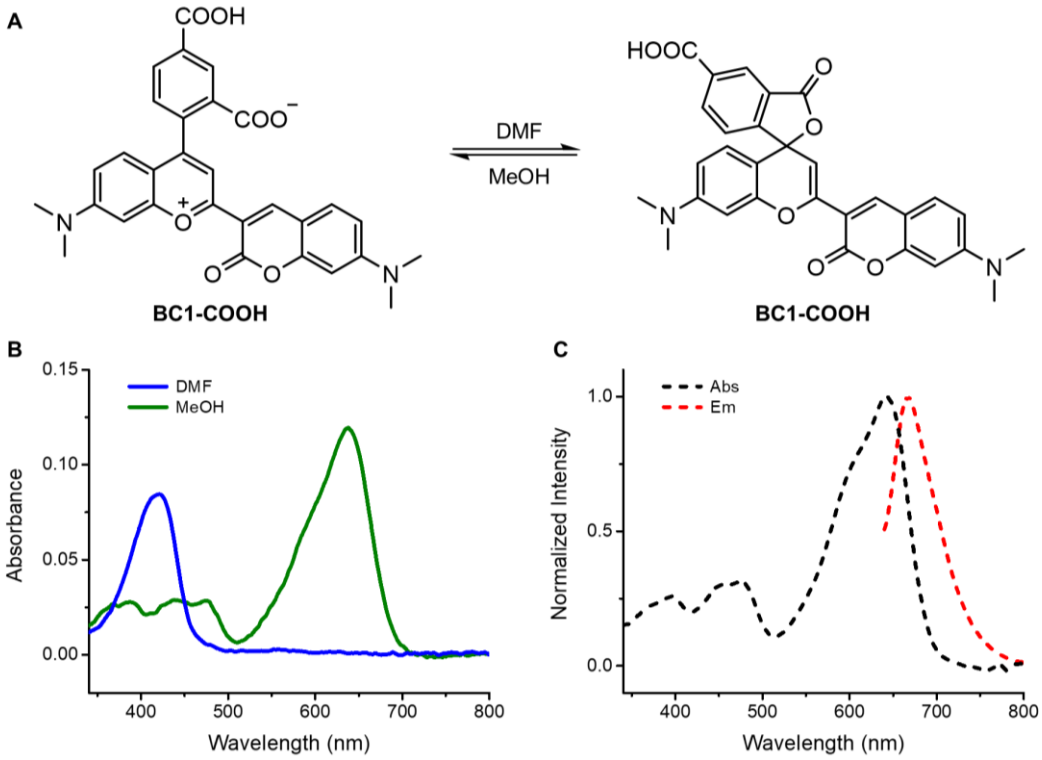
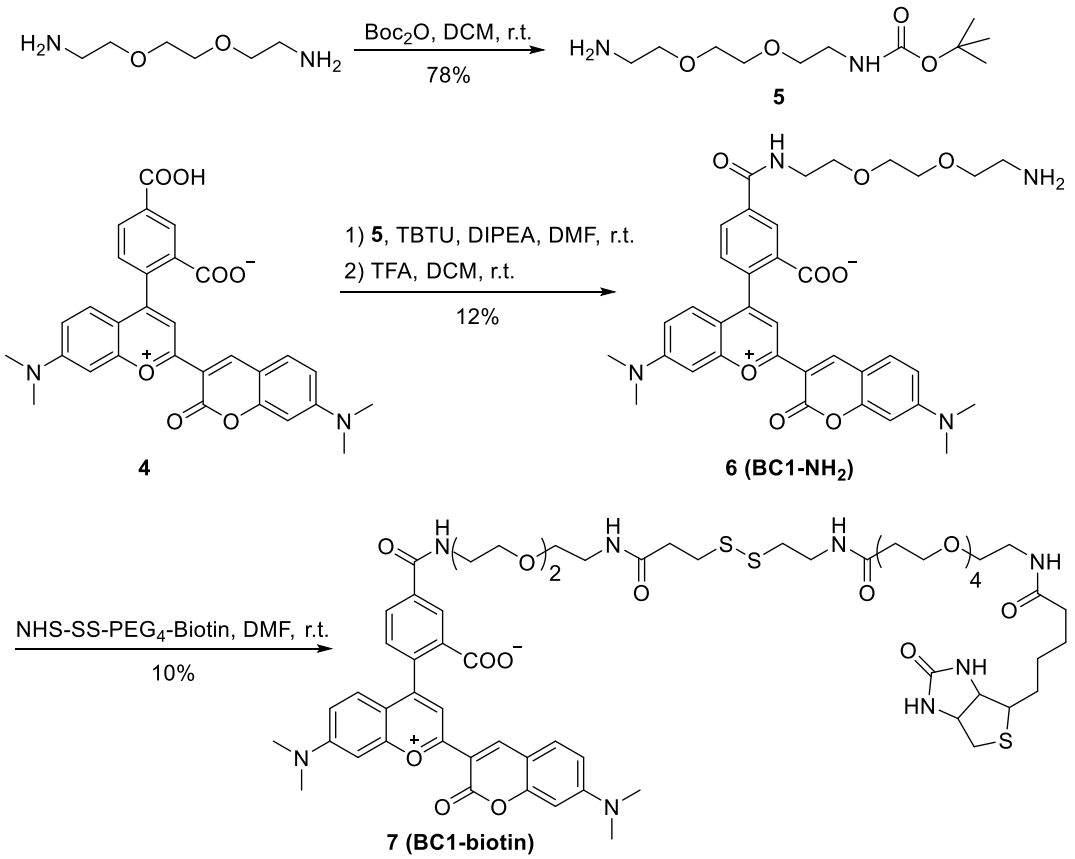


*Figure 25. Spectroscopic characterization of **BC1-COOH**. (A) Solvent-dependent intramolecular spirocyclization reaction of **BC1-COOH**. (B) Absorption spectra of **BC1-COOH** (2 µM) measured in DMF and MeOH. (C) The normalized absorption and emission spectra of **BC1-COOH** (1 µM) in 20 mM HEPES buffer. An excitation wavelength of 640 ± 5 nm was used.*

3.1.1.3 Synthesis of BC1-NH_2 and BC1-biotin

To attach the fluorophore to solid supports, we introduced the functional moieties (amine- and biotin-) to **BC1-COOH** via a PEG linker (Scheme 2). Amine-functionalized ligand **BC1-NH_2** was synthesized. **5** was prepared by a mono-tertbutyloxycarbonyl (Boc) protection of amino-triethylene glycol with a yield of 78%. Then, **5** was coupled to the sterically less hindered 5'-carboxyl group of **BC1-COOH** followed by deprotection of Boc under the acidic condition to render **BC1-NH_2** (**6**) with a 12% yield. Analogously, we obtained the biotin-functionalized ligand **BC1-biotin** via a coupling reaction between the commercial NHS-activated biotin-PEG and **BC1-NH_2**. Notably, the NHS-PEG-biotin contains a cleavable disulfide bond, which is useful for selective removal of **BC1**-binding aptamers at the selection phase.



*Scheme 2. Synthesis of the amine- and biotin-functionalized **BC1**: **BC1-NH$_2$** and **BC1-biotin**.*

3.1.2 *In vitro* selection of BC zwitterion-binding aptamer

Aptamers, RNA sequences that specifically bind to target ligands, are typically isolated from combinatorial RNA libraries. *In vitro* aptamer selection against **BC1** was carried out to obtain the specific RNA binders for the NIR-fluorescent BC zwitterion.

3.1.2.1 Systematic evolution of ligands by exponential enrichment

SELEX is a combinatorial chemistry technique for isolating aptamers that specifically bind to a target ligand. A typical SELEX cycle consists of five steps (Figure 26). (I) Negative selection. The RNA library is incubated with the mock resin to get rid of all non-specifically binding RNAs. (II) Binding. After removing the non-specific binders, the flow-through is then incubated with the ligand-labeled resin to initiate RNA-ligand interactions. Incubation conditions are crucial for selecting target-bound sequences. (III) Washing and partitioning. The non-binding or low binding RNAs are removed by washing thoroughly. Better partitioning results in fewer required selection cycles. (IV) Elution. The ligand-binding RNAs are subsequently eluted with elution buffer. Elution should be complete and specific. (V) Reverse transcription, PCR, and *in vitro* transcription. The eluted RNAs are reverse transcribed, PCR amplified, and *in vitro* transcribed.

Over-amplification should be avoided because it might introduce bias to the enriched sequences. The obtained RNA pool serves as input for the next round of selection. Generally, after 5-15 iterative rounds, the target-binding sequences are enriched and revealed by sequencing[103].

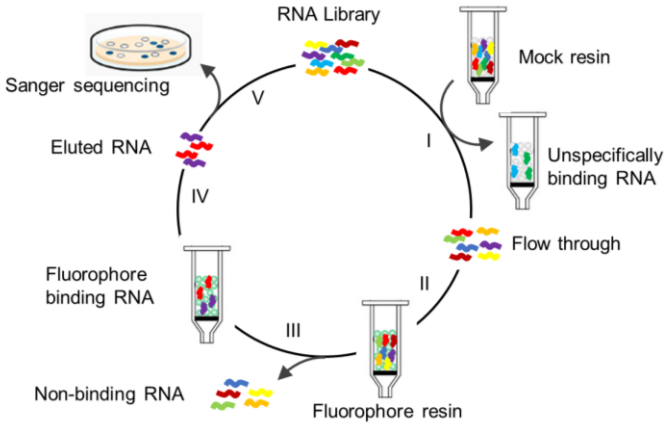


Figure 26. Scheme showing a typical SELEX cycle. I-Negative selection, II-Binding, III-Washing and partitioning, IV-Elution, V-Reverse transcription, PCR, and in vitro transcription.

3.1.2.2 Preparation of selection beads

BC1-NH$_2$ and **BC1-biotin** were attached to the solid supports in order to conduct the aptamer selection (Figure 27). First, **BC1-NH$_2$** was conjugated to NHS-activated sepharose beads in HEPES buffer through a standard amide coupling reaction to generate fluorophore-labeled resin for selection phase I. The **BC1** concentration (~mM) on beads was determined by subtracting the quantity of the unreacted dye from the input dye. In parallel, the mock resin was prepared by quenching all NHS groups on the sepharose beads with excess ethanolamine. To precisely control the ligand concentration on beads down to a µM level, streptavidin agarose beads were used for selection phase II. The strong biotin-streptavidin interaction ($K_D \approx 10^{-14}$ M) enabled the non-covalent labeling of **BC1-biotin** to the streptavidin beads after vigorous shaking in HEPES buffer[104]. In this way, the ligand concentration on beads could be determined by dividing the quantity of **BC1-biotin** by the volume of the beads. The streptavidin resin without any functionalization was used for mock selection.

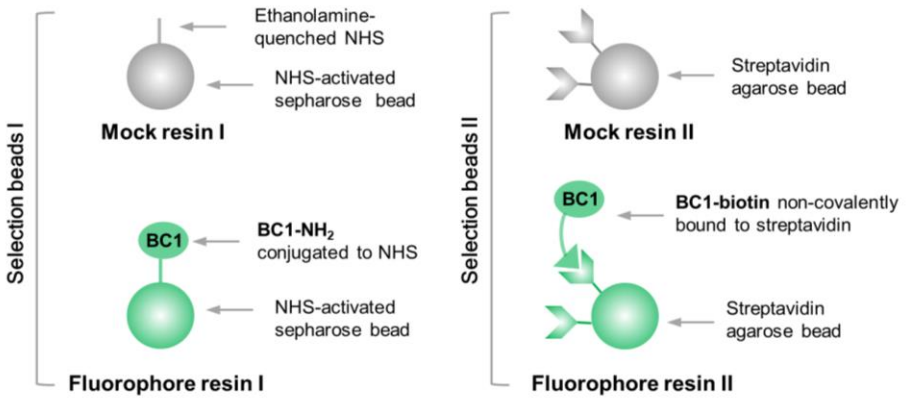


Figure 27. The resins as solid supports for the aptamer selection. The illustration of mock and fluorophore resins used for selection phases I and II is shown.

3.1.2.3 Design and preparation of RNA library

The combinatorial RNA library for aptamer selection was adapted from the reported literature[63, 105]. The RNA library (103-nt) comprises two fixed primer binding sites (19-nt and 20-nt, respectively), two 26-nt random regions, and a 12-nt constant region (Figure 28). The fixed primer-binding regions were designed for reverse transcription and polymerase chain reaction (PCR) amplification. The two random regions contributed to the library diversity. Moreover, the centrally-located stable stem-loop enabled a partially structured RNA pool, which was reported to generate more affinity binders than a completely random RNA library[105].

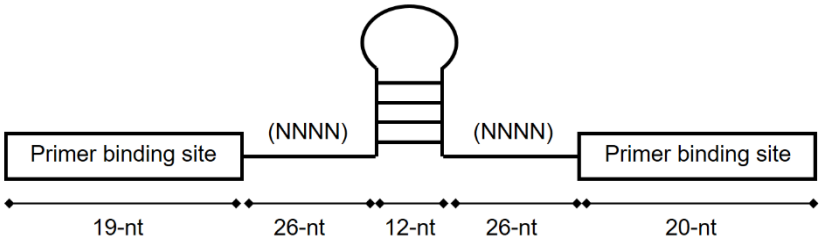


Figure 28. Design of the partially structured RNA library.

A single-stranded (ss) DNA library was chemically synthesized. RNA was transcribed from the DNA library on a massive scale (10 mL) and used as input for the aptamer selection (Table 2).

RNA library	5′ - GGA GCU CAG CCU UCA CUG C N_{26} CUG CUU CGG CAG N_{26} GGC ACC ACG GUC GGA UCC AC – 3′

Table 2. The sequence of the RNA library.

3.1.2.4 BC zwitterion-binding aptamer selection through SELEX

We started the selection with an RNA pool comprising a complexity of ~ 3×10^{15}, which is considered sufficient for successful aptamer isolation while retaining

simplicity during chemical synthesis[103]. Since affinity sequences were rare in the initial RNA pool, we started the selection with a high ligand concentration of 4 µM. This enabled an increased probability of interactions between the ligands and the RNAs. Mild washing conditions limited to 6 column volumes and harsh eluting conditions with elevated temperature were used to collect all the target-binding sequences at the first selection round. An elution buffer containing 98% formamide and 10 mM ethylenediaminetetraacetic acid (EDTA) was applied to deconstruct the aptamer folding and release the RNAs from the ligands[47, 106]. All these measures were set to increase the chances to fish out the **BC1**-binding sequences. The selection pressure was applied and increased progressively from round 2 by lowering the input RNA concentration from 100 µM to 10 µM and increasing the wash column volumes from 6 to 12. After the first five rounds (selection phase I), the RNA eluted from the fluorophore-labeled resin was increased from 0.06% to 35.3% of total input RNA, while this number for the negative selections was constantly around 0.5%, indicating the successful enrichment of **BC1**-binding aptamer (Figure 29). Notably, ligand concentration is an essential parameter for aptamer selection. The probability of isolating high-affinity binders increases with a lower ligand concentration. Therefore, we switched to streptavidin beads from the selection round 6 (selection phase II). The ligand concentration was decreased and carefully controlled by the use of **BC1-biotin**. Another six rounds of selection were performed. The selection stringency was further increased by declining the ligand concentration from 100 µM to 10 µM and increasing the washes to 30 column volumes. Furthermore, the elution buffer was switched to dithiothreitol (DTT) solution, which selectively cleaves the disulfide bond in the linker and releases the aptamer-fluorophore complex from solid supports. The SELEX ended at round 11 with an RNA elution percentage of 3.3%.

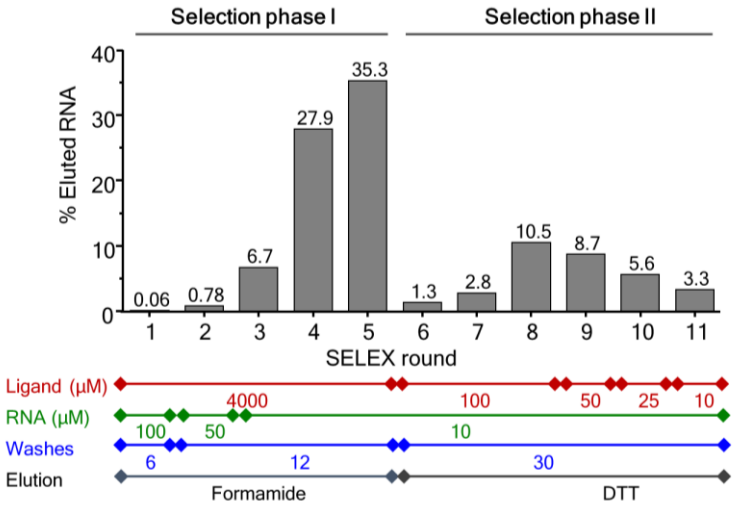


*Figure 29. Progress of the in vitro aptamer selection monitored by measuring the fraction of RNA eluted from the **BC1**-conjugated beads.*

After the SELEX, we examined the fluorescence performance of the RNAs eluted from each selection round towards **BC1-COOH** (Figure 30). Interestingly, the fluorescence of **BC1-COOH** increased slightly from round 1 to round 5. We speculated that the enhanced fluorescence signal was derived from the restriction of the non-radiative decay through twisted internal charge transfer (TICT) of the N,N-dimethylamino group of **BC1-COOH** by complexing with RNA[107]. However, the fluorescence started to decrease from round 6. This indicated that the selected higher affinity binders reduced the brightness of **BC1-COOH**, presumably due to fluorescence quenching by the spatially proximal RNA bases, especially guanosines to **BC1-COOH**[108-110].

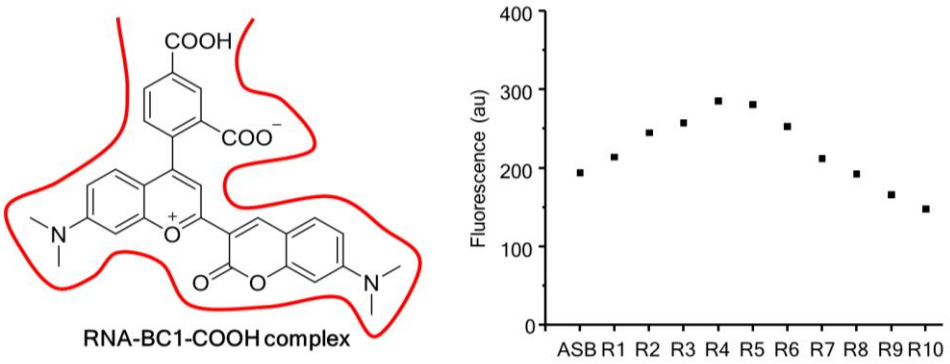


*Figure 30. The fluorescence performance of **BC1-COOH** (100 nM) in the presence of RNA (10 μM) from each selection round. The illustration of **BC1-COOH** complexed with an RNA aptamer (red curve) is shown. An excitation wavelength of 610 ± 5 nm and an aptamer selection buffer (ASB) containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl were used.*

The output RNA pool from the 11th round was reverse transcribed, cloned into a pDisplay vector, and transformed into *Escherichia coli* (*E. coli*). Ninety-six colonies were picked and sent for Sanger sequencing to determine the sequences of the potential aptamer candidates. Usually, the most abundant species or representatives of the enriched pool are believed to hold the affinity potential[103].

3.1.3 Identification of BC zwitterion-binding motif

The enriched RNA sequences from the SELEX were analyzed to identify the BC zwitterion-binding motif. Both binding affinity and brightness were taken into account when assessing the RNA sequences.

3.1.3.1 Sanger sequencing result

The Sanger sequencing results revealed 60 DNA sequences. Among them, six sequences were found twice, and one sequence was found three times (Figure 31). The 52 unique sequences were then subjected to multiple sequence alignments to identify common motifs. Most of them preserved the predefined tetraloop structure from the library design. However, it was difficult to discover

an evolved conserved motif between the sequences due to the large diversity of the sequencing pool even after using several different software.

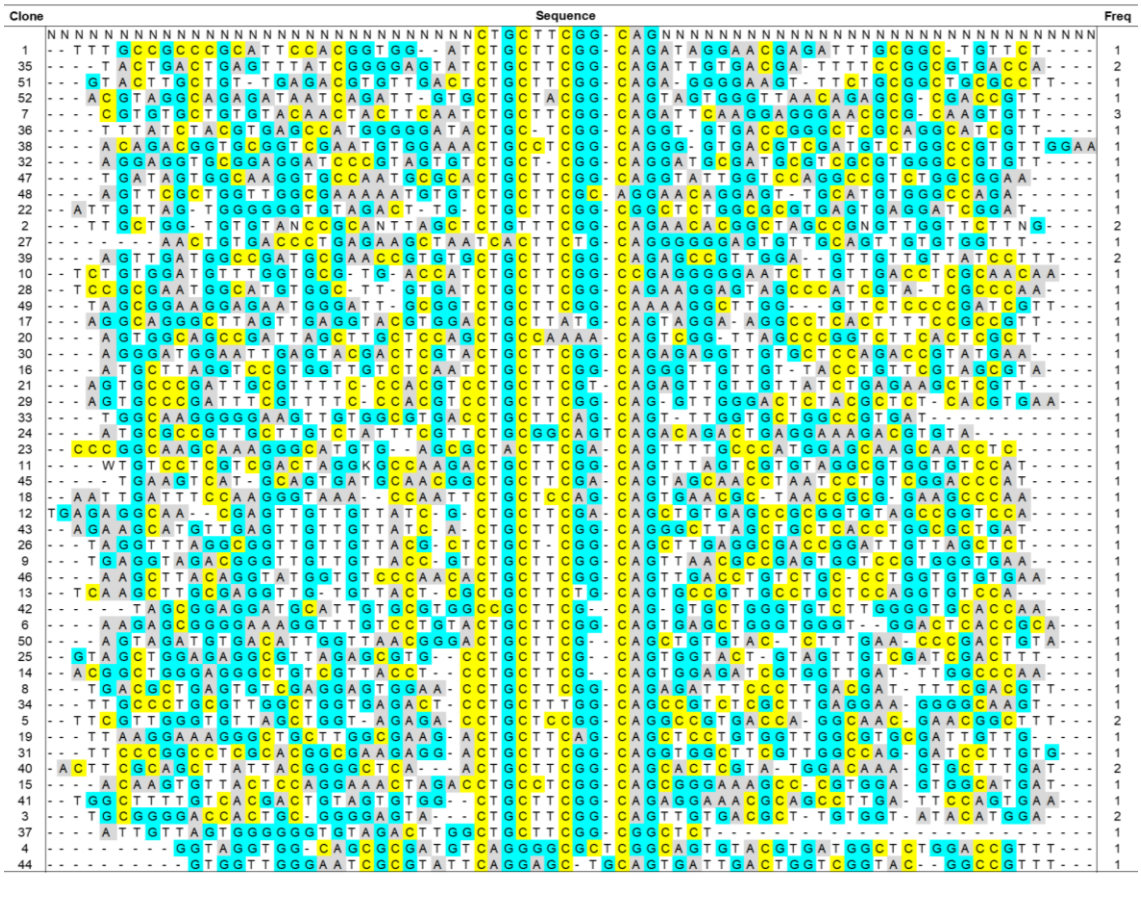


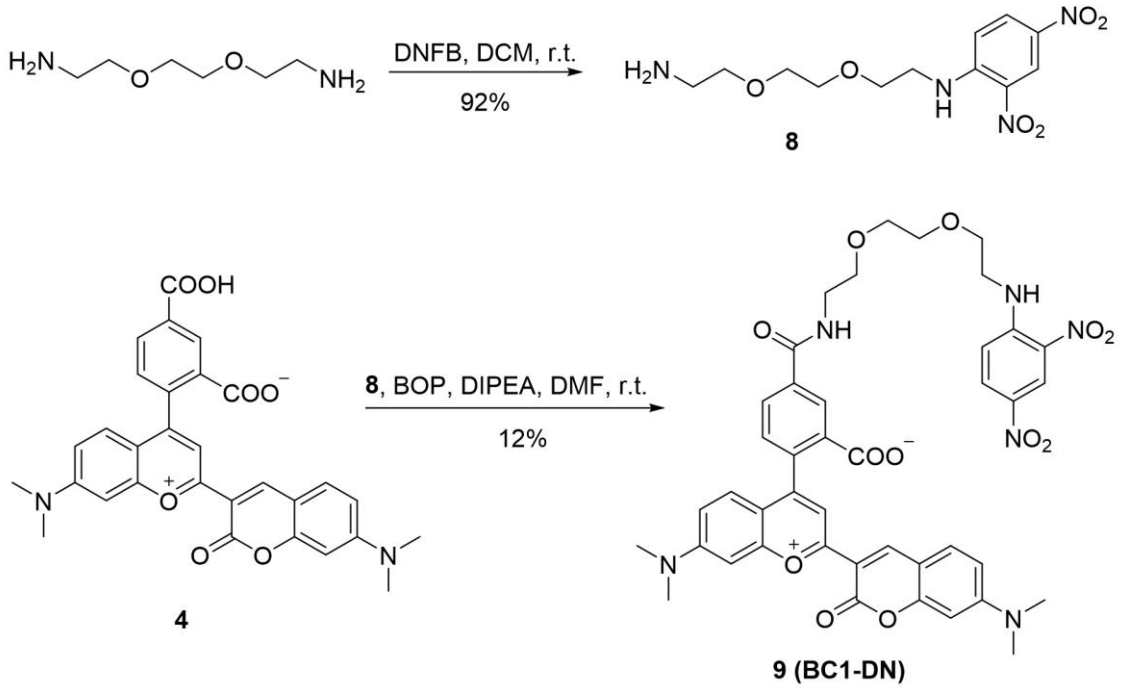
Figure 31. Sanger sequencing results of the selected clones and their frequencies. The primer-binding sequences at the 5' and 3' ends of each clone are not shown. The sequence alignment was performed using MultAlin (Symbol comparison table: blosum62; Gap weight: 12; Gap length weight: 2)[111].

3.1.3.2 Synthesis of BC1-DN

Using the zwitterionic **BC1** as a bait, RNA sequences that bind to **BC1** were successfully enriched through SELEX. The binding of **BC1** itself to the aptamer would not necessarily cause a fluorescence change. Therefore, a fluorogenic probe was required for evaluating the binding affinity and brightness of the selected RNA sequences to **BC1**. DN is a known contact quencher and has shown good quenching ability towards sulforhodamine B[56]. We introduced the DN to **BC1**, anticipating the formation of a non-fluorescent intramolecular complex **BC1-DN**. In the presence of a **BC1**-binding aptamer, **BC1** would prefer to interact with the aptamer rather than with the quencher, thereby inducing fluorescence enhancement.

BC1-DN was prepared according to the previously reported method (Scheme 3)[56]. 1-Fluoro-2,4-dinitrobenzene (DNFB) served as an electrophilic in the nucleophilic aromatic substitution (S_NAr) reaction with 2,2'-(ethylenedioxy)diethylamine, offering **8** with a 92% yield. This DN-PEG_3-NH_2

linker (**8**) was attached to **4** through an amide coupling to generate the contact-quenched **BC1-DN** (**9**).



*Scheme 3. Synthesis of **BC1-DN** (**9**).*

3.1.3.3 Activity screening

In vitro transcription and purification[a] rendered the 52 full-length RNAs (~ 103-nt) containing the forward and reverse primer sequences at 5′- and 3′- ends, respectively. The RNAs were incubated with **BC1-DN**, and the emission spectra were recorded to assess the fluorescence performance of the selected RNAs (Figure 32A). Out of 52 sequences, eight were determined to be highly active, showing more than 6-fold fluorescence enhancement ($F_{RNA\text{-}\mathbf{BC1\text{-}DN}}/F_{\mathbf{BC1\text{-}DN}}$), 29 sequences were moderately active with 3- to 6-fold fluorescence turn-on, and the remaining 15 were less active, possessing less than 3-fold fluorescence increment (Figure 32B). Besides, the emission maxima of the RNA-fluorophore complexes (λ_{em}: 687 - 715 nm) varied from that of **BC1-DN** (λ_{em} = 687 nm) in buffer solution, prominent evidence of the intermolecular interactions between aptamers and ligands (Figure 32C).

[a] The 52 RNA sequences were purified by phenol/chloroform/isoamyl alcohol extraction.

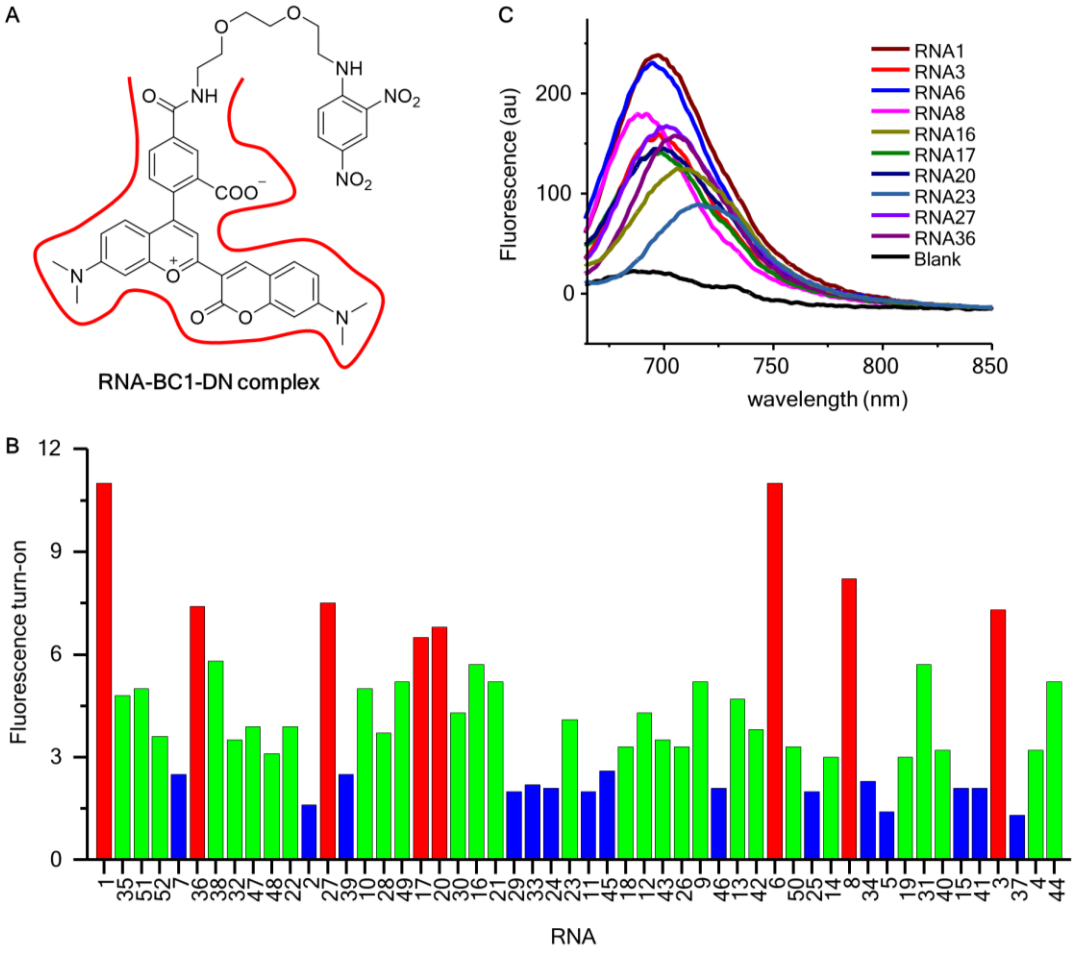


Figure 32. Fluorescence activity screening of the selected RNA sequences. (A) The illustration of ***BC1-DN*** *complexed with an RNA aptamer (red curve). (B) Fluorescence activity screening results of the selected RNA sequences. The RNAs (10 μM) were incubated with* ***BC1-DN*** *(100 nM), and the emission spectra were recorded with an excitation wavelength of 652 ± 5 nm. 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl was used. The fluorescence turn-on factors were determined by dividing the maximal fluorescence intensities of* ***BC1-DN*** *in the presence of RNA aptamers to that of* ***BC1-DN****. The highly (> 6-fold), moderate (3- to 6-fold), and less (< 3-fold) active RNAs are indicated in red, green, and blue, respectively. (C) Representative emission spectra of RNA sequences in the presence of* ***BC1-DN****.*

Brightness and binding affinity are two key parameters in assessing the value of a FLAP. The eight highly active sequences were further characterized with their binding affinities towards **BC1-DN** *in vitro*. We titrated **BC1-DN** with varying amounts of RNA and plotted the recorded maximal fluorescence intensity against RNA concentration (Figure 33). The eight RNAs displayed low-micromolar dissociate constants (K_D: 0.9 - 9 μM) and moderate fluorescence enhancement (turn-on: 9 - 16 fold). Among them, RNA8 showed the best binding affinity (K_D = 920 nM) and a promising fluorescence enhancement (10-fold) and was therefore selected as the aptamer candidate for further analysis.

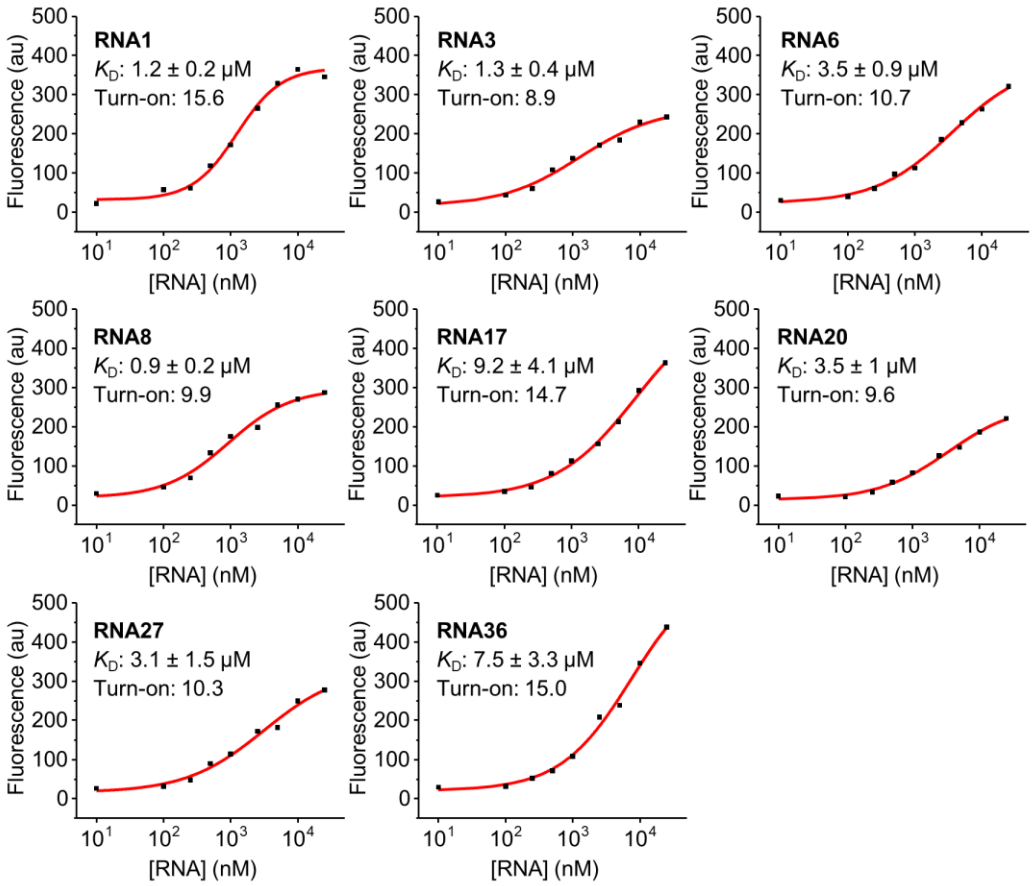


Figure 33. Characterization of the eight highly active RNA sequences. ***BC1-DN*** *(100 nM) was titrated with different amounts of RNA (0 - 25 μM) for each sequence individually. The K_D values were determined using Hill1 fit. The fluorescence turn-on factors were determined by dividing the maximal fluorescence intensities of* ***BC1-DN*** *in the presence of RNA (25 μM) by that of* ***BC1-DN****. For the fluorescence measurements, an excitation wavelength of 652 ± 5 nm and 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl were used.*

3.1.3.4 Truncation

Post-SELEX truncation was executed to minimize the aptamer size because shorter aptamers are preferred for intracellular FLAP imaging applications[43, 65]. Studies have shown that the primer-binding sequences are often minimally involved in the binding core[65, 112-113]. Exercising them could evolve a simpler aptamer structure with the same or even better affinity to the target. Therefore, we truncated the aptamer by deleting the primer-binding sites from both ends while preserving the two Gs at 5'-end, increasing the transcription efficiency by T7 polymerase (Figure 34). The truncated mutant RNA8-1 retained similar binding affinity (K_D = 0.9 μM) and fluorescence turn-on (9.1-fold) compared to the parental sequence RNA8. RNA8-2, with a further reduction of stem-loops based on the predicted secondary structure, preserved the intact fluorogenicity (9.7-fold) with an improved binding affinity (K_D = 0.2 μM). The increased affinity of RNA8-2 might derive from the shortening of the flanking sequences at both ends and the strengthening of the terminal stem by introducing a G-C pair, which

in total lowered the possibility of alternative-folding. However, an additional mutant RNA8-3, which lacks the internal hairpin, was less capable of inducing the fluorescence of **BC1-DN**, indicating the necessity of this centrally located stem-loop structure.

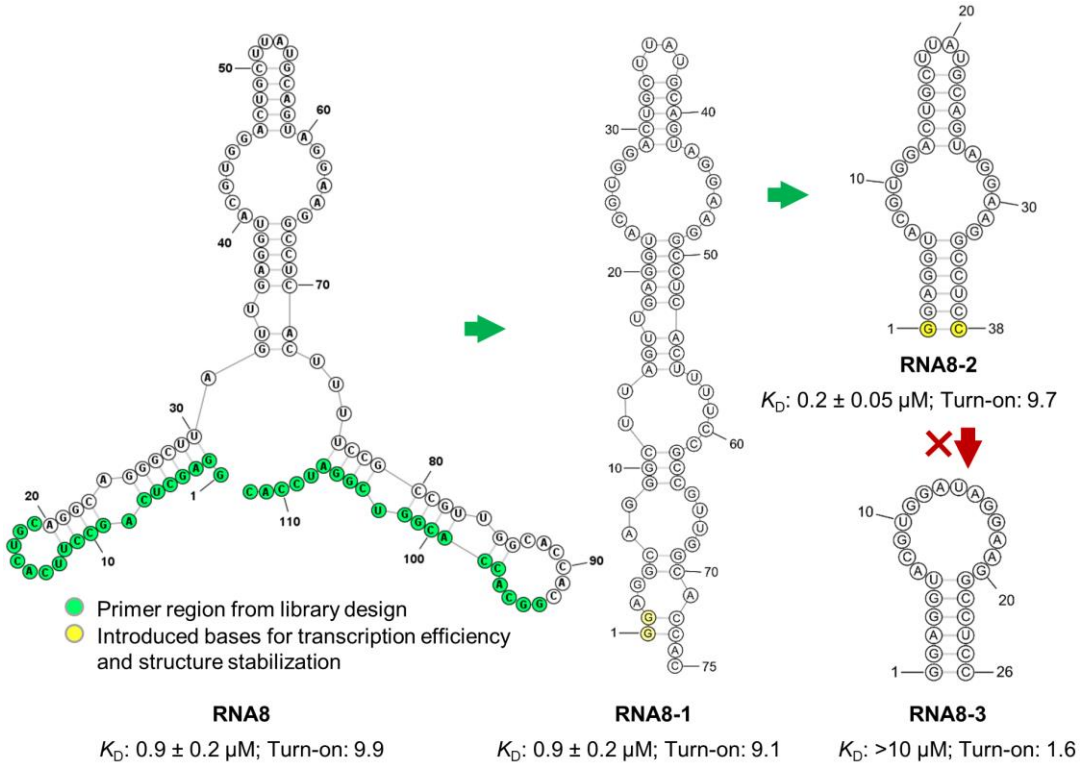


*Figure 34. Truncation analysis of RNA8. The predicted secondary structures of RNA8 and its truncated versions, RNA8-1, RNA8-2, RNA8-3, are shown. The K_D value was determined for each variant individually by titrating **BC1-DN** (100 nM) with different amounts of RNA (0 - 10 μM). The fluorescence turn-on factors were determined by dividing the maximal fluorescence intensities of **BC1-DN** (100 nM) in the presence of RNA mutants (10 μM) by that of **BC1-DN**. For the fluorescence measurements, an excitation wavelength of 652 ± 5 nm and 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl were used.*

3.1.3.5 Common motif identification

RNA8-2 was identified as the core sequence for the fluorophore binding. Interestingly, re-analysis of the eight highly active sequences and their predicted secondary structures revealed that RNA1 and RNA20 had similar core structures as RNA8 (Figure 35). Except for the terminal stem-loop from the library design, the yellow-labeled nucleotides GUGG and AGGAA in the main loop were conserved in all three sequences, denoting their crucial roles in fluorophore complexation. This finding demonstrated that the main loop of RNA8-2 was a common motif evolved through SELEX and responsible for the high-affinity binding. The three RNA sequences' homology was not discovered at the earlier stage when sequence alignment was conducted, presumably due to the small conservated region (8-nt) out of the full-length (64-nt) that was not strong enough for the software to sort them into the same cluster.

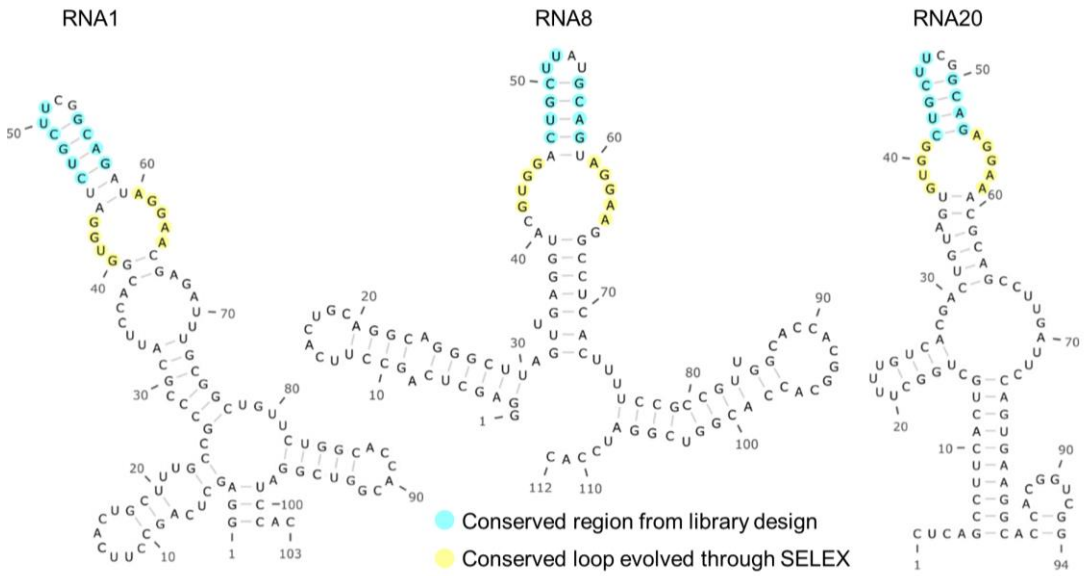


Figure 35. Common motif identification. Predicted secondary structures of RNA1, RNA8, and RNA20 are shown. The conserved nucleotides evolved through SELEX and the conserved nucleotides from the library design are highlighted in yellow and blue, respectively.

Moreover, we found that many of the 52 isolated RNAs possessed a loop that was the flipped or rotated variant of the main loop of RNA8-2. Two RNA variants possessing either a horizontal flipped or a 180° rotated loop relative to that of RNA8-2 were designed to investigate the spatial arrangement of the 12-nt loop and its influence on RNA-fluorophore complexation (Figure 36A). Surprisingly, the original, flipped, and rotated RNA8-2 showed similar fluorescence enhancement towards **BC1-DN** under a high aptamer concentration despite their different structural orientations (Figure 36B). Only under a low aptamer concentration, the original RNA8-2 exhibited a slightly higher fluorogenicity, indicating a better binding affinity between RNA8-2 and **BC1-DN**. Although the 3-dimensional structure of RNA8-2 remains unknown, this observation suggested that the main loop is the binding core. The two helices are responsible for the loop generation and are minimally involved in the fluorophore complexation since the changes in the loop's structural orientation did not significantly abolish the binding.

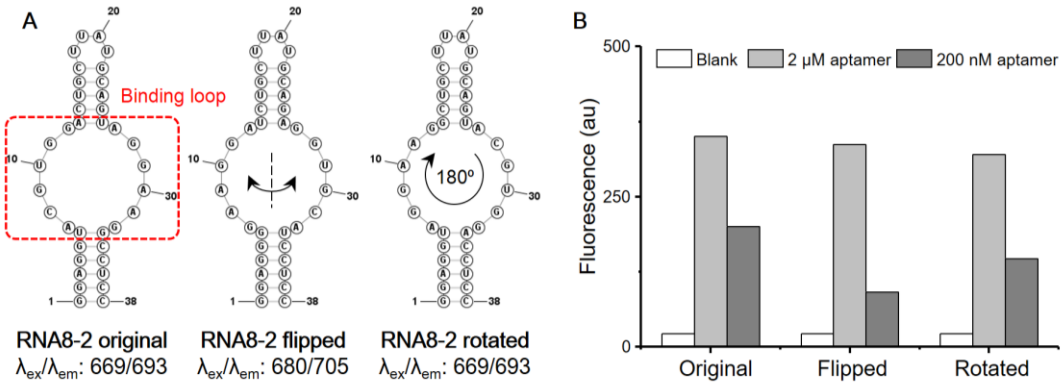


*Figure 36. Fluorescence performance of the original, flipped, and rotated RNA8-2. (A) The predicted secondary structures of original, flipped, and rotated RNA8-2. The excitation and emission maxima of the RNA-**BC1-DN** complexes are shown in nm. (B) Fluorescence activity screening of original, flipped, and rotated RNA8-2. **BC1-DN** (100 nM) was incubated with no, low (200 nM), and high (2 μM) concentrations of RNA. The maximum fluorescence intensities were recorded at their excitation maxima (± 5 nm). 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl was used.*

3.1.3.6 Mutation

A total of 12 mutants were designed to assess the effects of the non-conversed nucleotides in the proposed motif of RNA8-2 (Figure 37). The respective RNAs were produced and assayed for both fluorescence enhancement and binding affinity using fluorogenic **BC1-DN**.

RNA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
8-2	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	C	C	U	C	C
M1	G	G	A	G	G	C	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	C	C	U	C	C
M2	G	G	A	G	G	G	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	C	C	C	U	C	C
M3	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	A	C	C	U	C	C
M4	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	C	C	C	U	C	C
M5	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	U	C	C	U	C	C
M6	G	G	A	A	A	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	U	U	U	C	C
M7	G	G	A	G	-	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	-	C	U	C	C
M8	G	G	A	G	G	U	-	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	C	C	U	C	C
M9	G	G	A	G	G	U	A	-	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	G	C	C	U	C	C
M10	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	A	U	G	C	A	G	U	A	G	G	A	A	G	-	C	C	U	C	C
M11	G	G	A	G	G	U	A	C	G	U	G	G	G	C	U	G	C	U	U	A	U	G	C	A	G	C	A	G	G	A	A	G	G	C	C	U	C	C
M12	G	G	A	G	G	U	A	C	G	U	G	G	A	C	U	G	C	U	U	C	G	G	C	A	G	U	A	G	G	A	A	G	G	C	C	U	C	C

Figure 37. The sequences of RNA8-2 and the 12 mutants. The mutation sites are indicated in green.

We first tried to "fix" the wobble pair U_6-G_{33} in the bottom helix since it is not a conventional Watson-Crick base pair. We found that the U_6-G_{33} pair was crucial since exchanging it to C-G (M1), G-C (M2), U-A(M3), U-C(M4), U-U(M5) significantly attenuated fluorophore binding as indicated by the increased K_D values (Figure 38). Next, we evaluated the influence of the terminal helix's length and stringency on the fluorophore binding. However, exchanging G_4G_5-$C_{34}C_{35}$ to AA-UU (M6) or deleting base-pair G_5-C_{34} (M7) showed negligible effects. Moreover, we checked the effects of A_7, C_8, G_{33} in the main loop by deleting them individually (M8, M9, M10, respectively). All three mutants showed decreased fluorescence turn-ons and worsened dissociate constants, demonstrating the critical roles of the three nucleotides in the aptamer-fluorophore complexation. Lastly, we exchanged the A_{13}-U_{26} pair to a stronger base pair G-C (M11) or replaced the $A_{20}U_{21}$ with the original CG (M12) in the tetraloop. Both of them failed to promote the performance of the aptamer further.

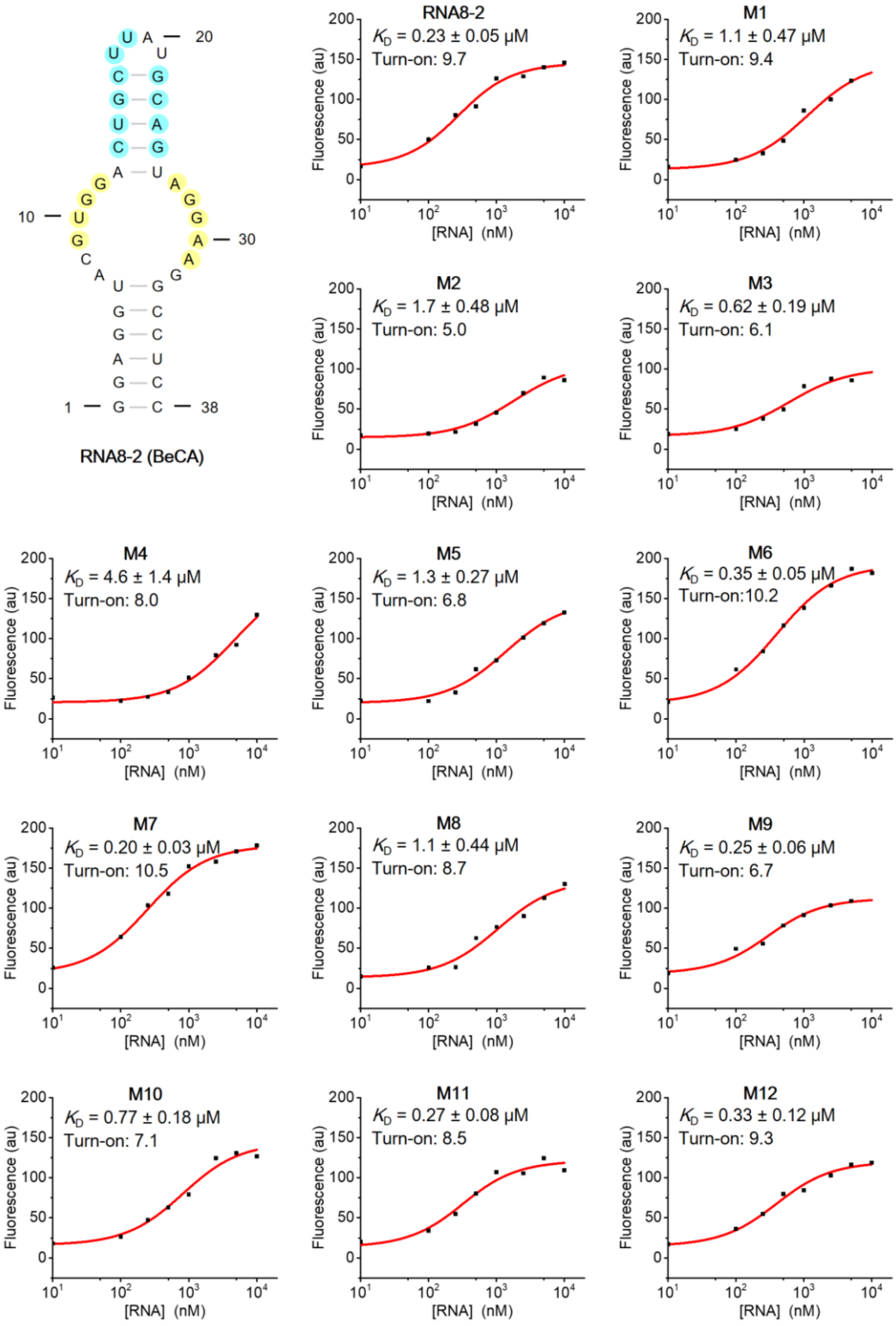


Figure 38. Activity screening of the RNA8-2 mutants. The proposed aptamer structure of RNA8-2 is shown. The K_D values of RNA8-2 and all mutants were determined by titrating ***BC1-DN*** *(100 nM) with varying concentrations of RNA (0 - 10 μM) individually. The fluorescence turn-on factors were determined by dividing the maximal fluorescence intensities of* ***BC1-DN*** *in the presence of RNA mutants (10 μM) by that of* ***BC1-DN****. An excitation of 652 ± 5 nm and 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl were used.*

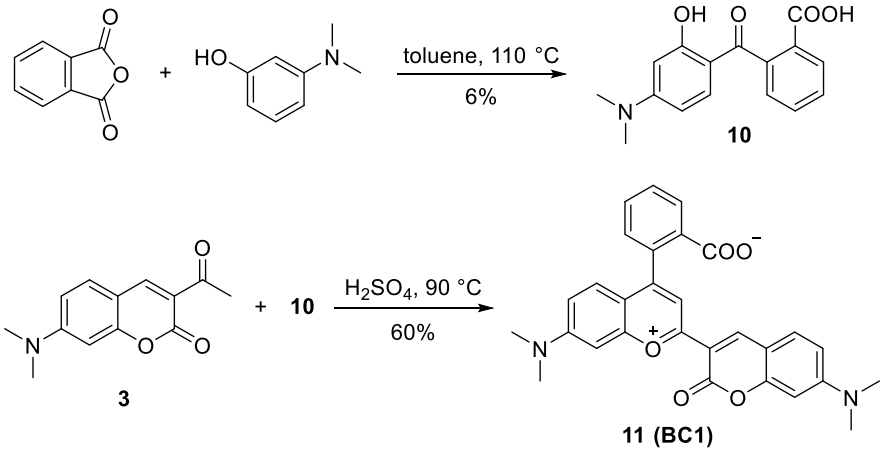
By taking all bioinformatic and fluorescence data together, we successfully identified a 38-nt benzopyrylium-coumarin binding aptamer RNA8-2 (hereafter referred to as **BeCA**) that binds to **BC1-DN** with nanomolar affinity and good fluorogenicity.

3.1.4 Synthesis and characterization of color-switchable benzopyrylium-coumarin analogs

The **BeCA** aptamer was designed and evolved to bind **BC1** zwitterion selectively. To obtain a color-shifting aptamer-fluorophore module, alternative ligands that stay in the cyan-fluorescent spirocyclic form in the unbound state but efficiently switch to the NIR-fluorescent zwitterion form upon binding to **BeCA** were needed.

3.1.4.1 Synthesis of BC1

The mono-carboxylic **BC1** was synthesized similarly to its dual-carboxylic counterpart **BC1-COOH** according to the reported literature (Scheme 4)[84]. First, the acylation of 3-dimethylaminophenol with phthalic anhydride gave benzoylbenzoic acid (**10**). The condensation between **10** and coumarin (**3**; chapter 3.1.1.1) rendered **BC1** with a 60% yield.



*Scheme 4. Synthesis of **BC1 (11)**[b].*

3.1.4.2 Characterization of BeCA-BC1

The SELEX was designed and performed to select aptamers binding to BC zwitterion, i.e., **BC1**. Nevertheless, NH_2-PEG and biotin-PEG linkers were appended to **BC1** for the ligand conjugation to solid supports and were potentially involved in the aptamer-ligand complexation. The subsequent removal of the linker from the substrate could potentially cause a loss of affinity[65,]

[b] The **BC1** was synthesized by Elena Papalozi during her internship in Jäschke laboratory under my supervision.

[114]. We examined the binding affinity between **BeCA** and **BC1** without any linker attachment (Figure 39). Fortunately, **BC1** displayed a K_D of 320 nM towards **BeCA**, slightly weaker than **BC1-DN** (K_D = 230 nM). Thus, we concluded that **BeCA** is an adequate binder for **BC1**. The observed fluorescence turn-on could be reasoned to the quantum yield increment of **BC1** either from the break of dye aggregation or restriction of TICT upon aptamer binding.

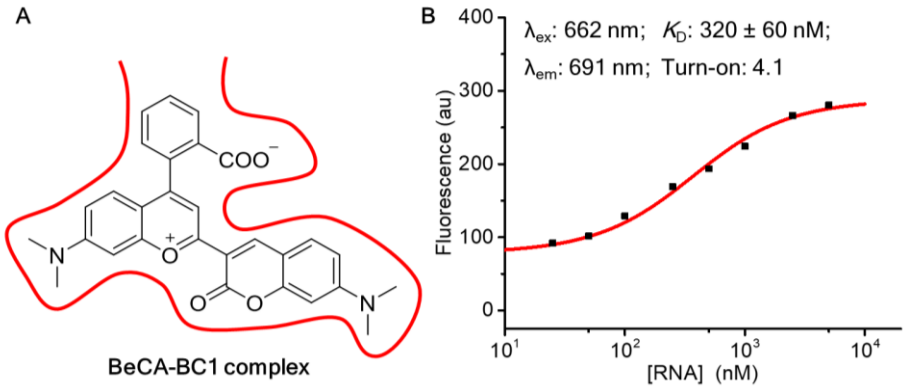


*Figure 39. Characterization of **BeCA-BC1** complex. (A) The illustration of **BC1** complexed with **BeCA** (red curve). (B) The K_D curve of **BeCA-BC1** complex. The K_D value was determined by titrating **BC1-DN** (100 nM) with different amounts of RNA (0 - 5 μM). An excitation of 662 ± 5 nm and 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl were used.*

3.1.4.3 Design of BC analogs

BC1, as the binding ligand for **BeCA**, favors the NIR-fluorescent zwitterionic form in aqueous solutions. BC fluorophores that change their emission colors upon **BeCA** binding were required to develop color-shifting FLAPs. We performed chemical modification on the **BC1** scaffold to alter the intramolecular equilibrium, promoting its propensity to form the cyan-fluorescent spirocyclic form. It was reported that increasing either the electrophilicity of the chromophore or the nucleophilicity of the intramolecular nucleophile could shift rhodamines' equilibrium to the spirocyclic configuration[80, 82, 95, 115]. Thus, two approaches were applied to tune the spirocyclization reaction of BC fluorophores (Figure 40). In the first approach, the electron-withdrawing fluorine atom was introduced to the chromophore at the aromatic positions of the benzopyrylium and coumarin moieties. In the second approach, electron-deficient amines were introduced to the carboxyl group of the upper benzene ring.

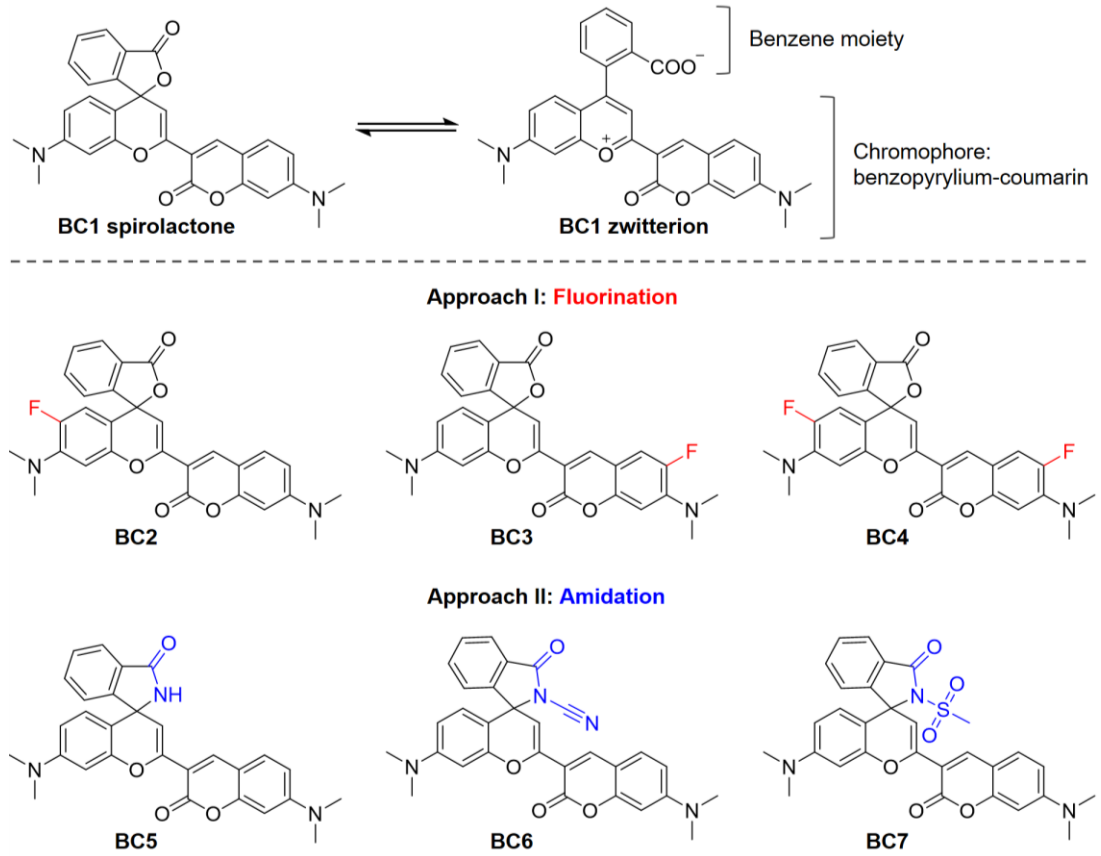
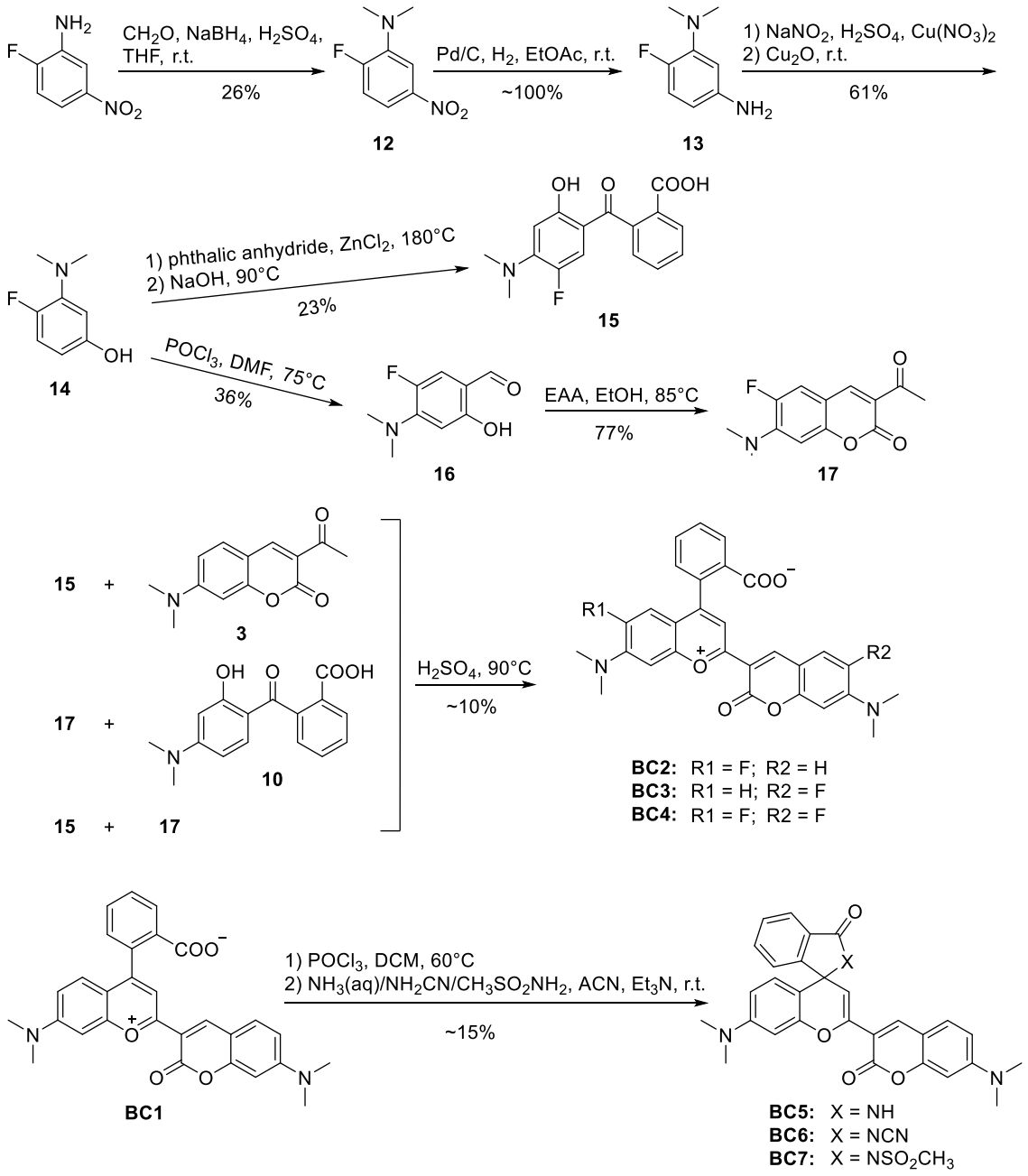


*Figure 40. Design of the spirocyclic BC analogs. The intramolecular equilibrium of **BC1** is shown. Chemical alterations of **BC1** on either the benzene ring or the chromophore generated **BC2-7**. The modification sites are indicated in red or blue according to the different design strategies, fluorination, or amidation of **BC1**.*

3.1.4.4 Synthesis of BC2-7

To generate the fluorinated **BC1**, we first synthesized 2-fluoro-5-hydroxy-N,N-dimethylaniline (**14**) according to the reported literature (Scheme 5)[116]. The reductive amination was applied to 2-fluoro-5-nitroaniline to install two methyl groups to the amine, providing **12** with a yield of 26%. Then, the hydrogenation of the nitro group of **12** in the presence of a palladium catalyst gave **13** with a quantitative conversion. **14** was obtained via a *Sandmeyer* reaction and the subsequent hydroxylation. The condensation between benzoylbenzoic acids and coumarin derivatives (**15-3**, **10-17**, or **15-17**) yielded **BC2-4,** respectively. Alternatively, **BC5-7** were achieved by the standard amide coupling. The carboxyl group of **BC1** was first activated to an acid chloride and then reacted with ammonia, cyanamide, and methanesulfonamide, respectively, to render **BC5-7**.



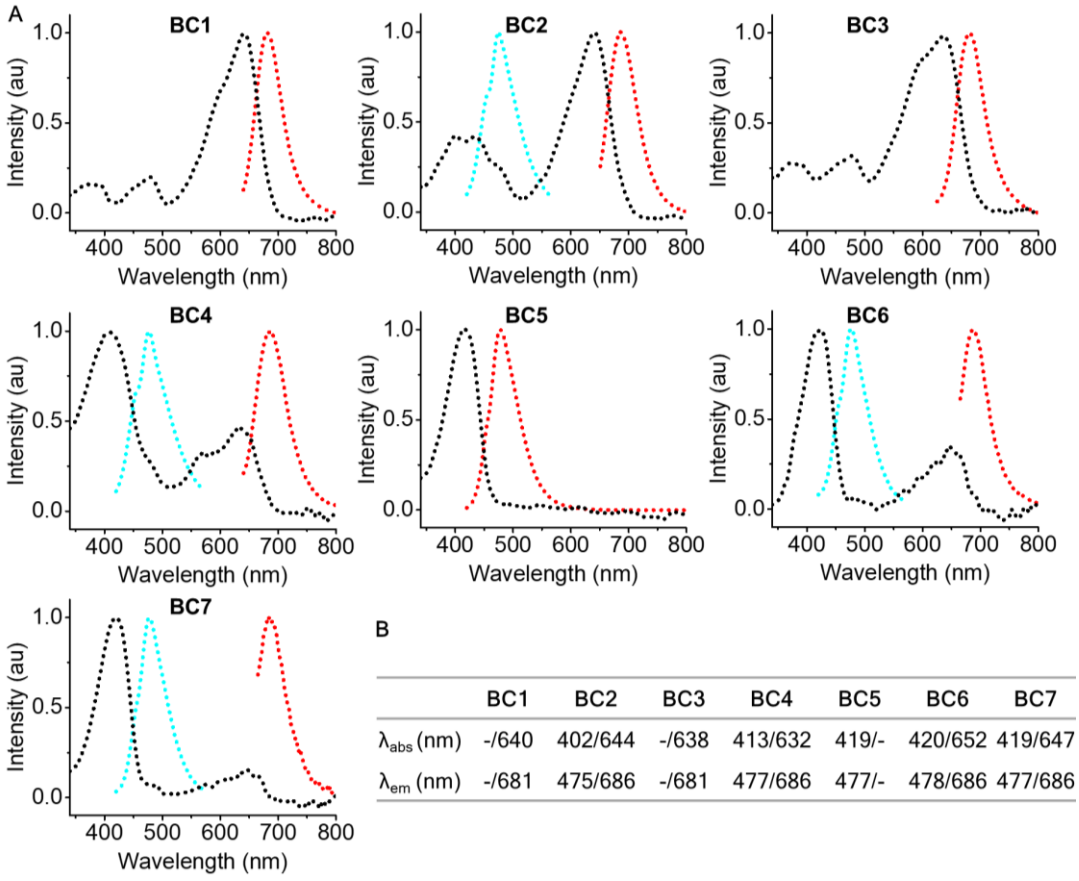
*Scheme 5. Synthesis of **BC2-7**[c].*

3.1.4.5 Spectroscopic characterization of BC1-7

The absorbance and emission spectra of **BC1-7** in HEPES buffer were collected (Figure 41). **BC1** and **BC3** showed absorption peaks only from the zwitterionic form (λ_{abs} = 640 and 638 nm, respectively) and fluorescent in the NIR region (λ_{em} = 681 nm for both). On the contrary, **BC5** exhibited a mono-absorption peak from

[c] The compounds **BC2-4** were synthesized by Leon Tydecks, Martina A. Pogány, and Paul Haubenwallner during their internships in Jäschke laboratory under my supervision.

the spirocyclic form (λ_{abs} = 419 nm) and emitted solely in the cyan window (λ_{em} = 477 nm). The rest BCs (**BC2**, **BC4**, **BC6**, and **BC7**) displayed absorption and emission from both spirocyclic and zwitterionic forms with varying intensity ratios. These observations demonstrated that BC analogs indeed possessed an intramolecular equilibrium between two configurations. The zwitterion/spiro-lactone (lactam) ratio in aqueous solutions was related to their chemical properties.



B

	BC1	BC2	BC3	BC4	BC5	BC6	BC7
λ_{abs} (nm)	-/640	402/644	-/638	413/632	419/-	420/652	419/647
λ_{em} (nm)	-/681	475/686	-/681	477/686	477/-	478/686	477/686

*Figure 41. Spectroscopic characterization of **BC1-7**. (A) Normalized absorbance and emission spectra of **BC1-7** (1 μM) in 20 mM HEPES containing 0.1% Triton X-100. Black, cyan, and red dotted lines represent the absorbance, emission spectra of the spirocyclic form, and the emission spectra of the zwitterionic form, respectively. (B) Table showing the absorbance and emission maxima of **BC1-7** in the format of spiro-lactone(lactam)/zwitterion. The data was generated from panel A.*

3.1.4.6 Dielectric constant titration of BC1-7

BC fluorophores possess an intramolecular equilibrium between a spirocyclic form and a zwitterionic form. To precisely evaluate how the chemical modification influences the lactone(lactam)-zwitterion equilibrium, dielectric constant titration of **BC1-7** was executed in the water/dioxane mixtures. Dioxane possesses a low dielectric constant (ε_r = 2.2) and favors the spirocyclic configuration of BC[73, 82]. On the contrary, water has a high dielectric constant (ε_r = 80.4) and promotes the formation of zwitterionic BC. D_{50} value is defined as the

dielectric constant at which half of the maximal absorbance of zwitterion was observed. A higher D_{50} value means a higher degree of spirocyclization. Absorption spectra of **BC1-7** in the dioxane-water mixtures with different ionic strengths were measured, and the D_{50} values were determined (Figure 42). The titration indicated that **BC1** existed predominantly in its zwitterionic form in the aqueous solution with a low D_{50} value of 23. The introduction of the fluorine atom to the benzopyrylium and coumarin moieties shifted the equilibrium to the spirocyclic form, indicated by the increased D_{50} values (39 for **BC2**, 30 for **BC3**, and 46 for **BC4**). Nevertheless, reacting the carboxyl group to a primary amide converted **BC5** completely to the spirocyclic form even in pure water, displaying a D_{50} value >> 70. Using electron-deficient amides alleviated the "ring-closing" process, resulting in reduced D_{50} values: 51 for **BC6** and 62 for **BC7**.
In conclusion, the chemical modifications on the **BC1** scaffold fine-tuned the equilibrium to the spirocyclic form demonstrated by the increased D_{50} values. Rhodamine-based fluorophores with D_{50} values around 50 are suitable candidates for the generation of cell-permeable and fluorogenic probes[82]. Therefore, we anticipated that **BC2**, **BC4**, **BC6**, **BC7** with D_{50} ranging from 40 to 60 held the potential to function as the color-shifting ligands for **BeCA**.

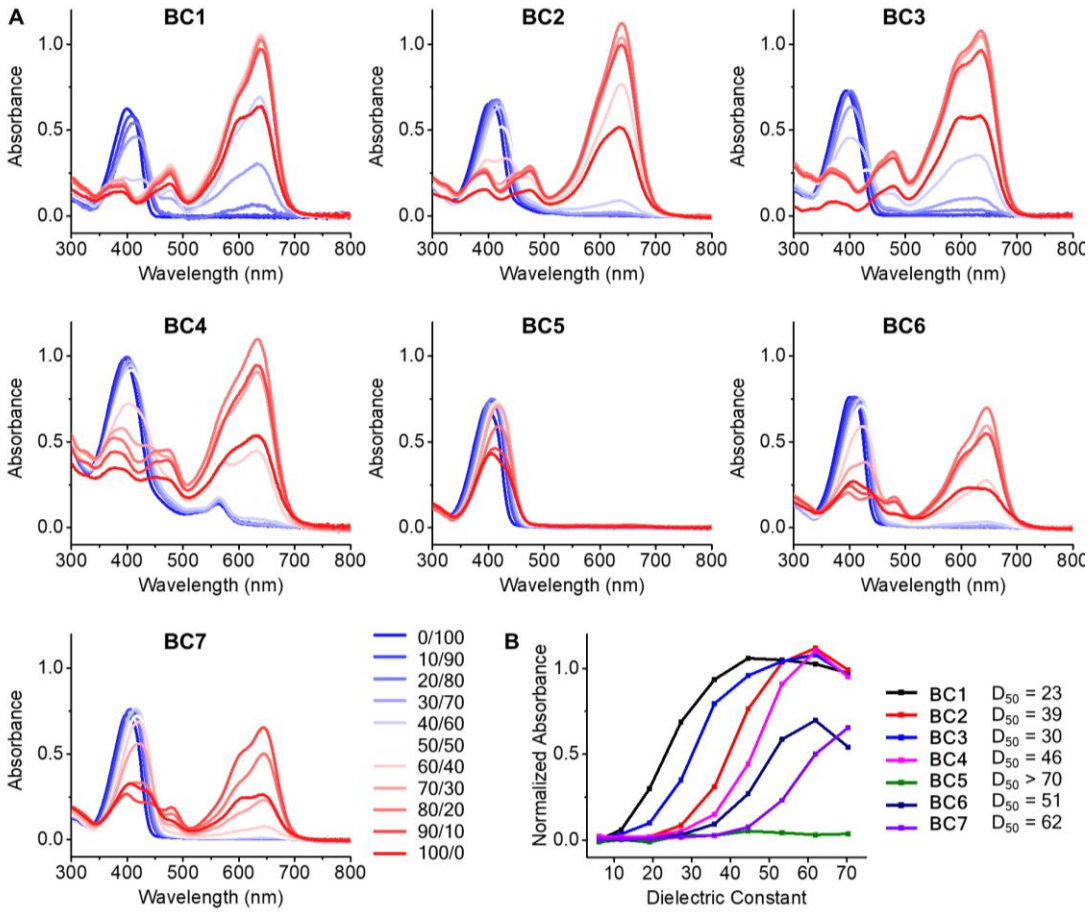


*Figure 42. Dielectric constant titration of **BC1-7**. (A) Normalized absorption spectra of **BC1-7** (5 μM) in water/dioxane mixtures (v/v: 0/100 - 100/0). The absorbance of **BC1-7** was normalized to that of each probe*

*in water containing 0.3% sodium dodecyl sulfate (SDS), in which the fluorophores were predominantly converted to their corresponding zwitterions. (B) Maximal zwitterionic absorbance of **BC1-7** as a function of dielectric constant. The data points were generated from panel A and plotted.*

3.1.5 Substrate screening for BeCA

Substrate screening was performed to evaluate the fluorescence performance of BC analogs towards **BeCA** binding. First, **BeCA** was incubated with **BC1-7** respectively, and the fluorescence turn-on in the NIR channel was recorded (Table 3). They exhibited fluorescence enhancement ranging from 1.6- to 8.5-fold. Among them, **BC2** and **BC6** showed the highest fluorescence turn-ons (8.5-fold and 8.2-fold, respectively), which were comparable to that of the contact-quenched probe **BC1-DN**. We reasoned that the fluorogenicity was derived from the break of dye aggregation, shift of equilibrium to the NIR-fluorescent zwitterion, or restriction of TICT upon **BeCA** binding.

	BC1	**BC2**	**BC3**	**BC4**	**BC5**	**BC6**	**BC7**
NIR fluorescence turn-on	3.9	8.5	2.7	2.1	1.6	8.2	2.1

*Table 3. Fluorescence performance of **BC1-7** towards **BeCA** binding. **BC1-7** (100 nM) were incubated without or with **BeCA** (10 μM) and excited at their excitation maxima (± 5 nm) in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl. The NIR turn-on factors were determined by dividing the maximal NIR fluorescence intensities of BC fluorophores in the presence of **BeCA** by those of BC fluorophores.*

Triton, a nonionic surfactant, prevents aggregation in aqueous solutions by interacting with the lipophilic dye[117]. We added 0.05% Triton X-100 into the measuring buffer to exclude any fluorescence enhancement derived from the disruption of aggregation-caused quenching (ACQ) of the fluorophore upon aptamer binding, thereby better evaluating the color-shifting ability of BC analogs due to the equilibrium shift. The absorption and emission spectra of **BC1-7** towards **BeCA** binding were recorded.

In theory, the affinity between **BeCA** and BC analogs could be potentially disturbed due to the chemical modifications on the **BC1** scaffold. However, all BC substrates except **BC5** showed a sharper and red-shifted absorption peak in the presence of **BeCA** aptamer, indicating the formation of the **BeCA**-BC complex (Figure 43). Among them, **BC2**, **BC4**, **BC6**, and **BC7** with D_{50} around 50 showed a decrease in the spirocyclic absorption peak ($\lambda_{spirocyclic\text{-}abs} \approx 420$ nm) and an increase in the zwitterionic absorption peak ($\lambda_{zwitterion\text{-}abs} \approx 680$ nm), demonstrating the equilibrium shift of the probes from the spirocyclic form to the zwitterionic form upon aptamer binding.

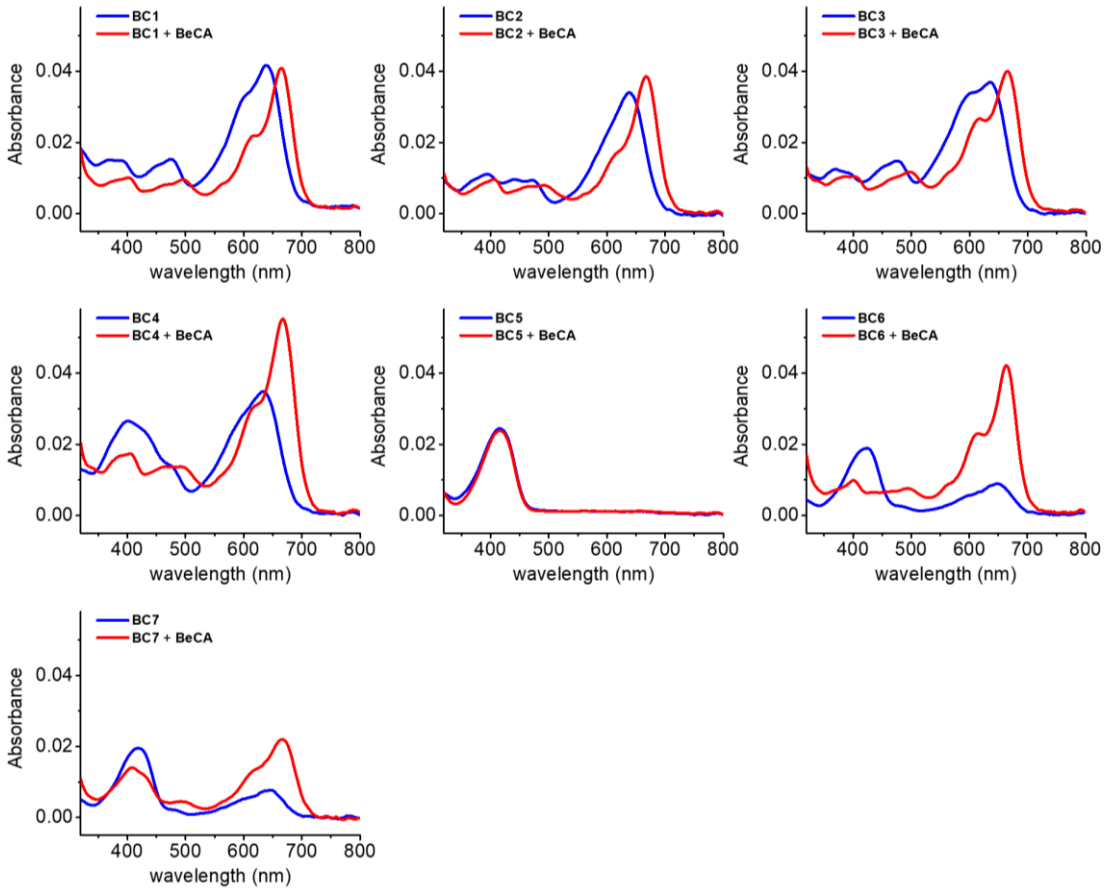
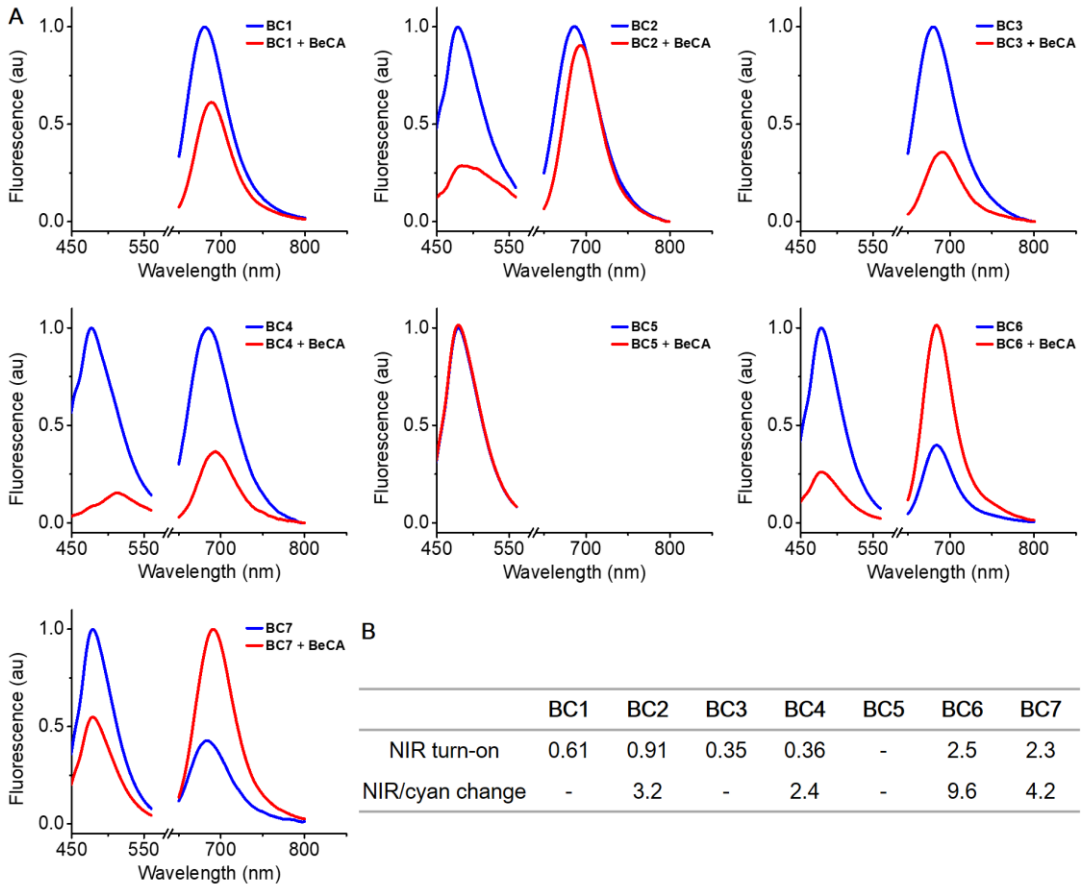


*Figure 43. Absorption spectra of **BC1-7** towards **BeCA** binding. **BC1-7** (1 μM) were incubated in the absence or presence of **BeCA** (25 μM) in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100.*

The spirocyclic **BC5** displayed no fluorescence change since it barely interacted with the BC zwitterion-binding aptamer **BeCA** (Figure 44A). The cyan fluorescence of **BC2**, **BC4**, **BC6**, and **BC7** declined upon the addition of **BeCA** due to the reduced amount of the unbound spirocyclic BC. Surprisingly, only **BC6** and **BC7** showed enhanced NIR fluorescence (2.5- and 2.3- fold, respectively) as a result of the formation of the **BeCA**-BC complex. **BC2** and **BC4** showed no rise or even a decline in the NIR channel, although the binding of their zwitterionic forms to **BeCA** was indicated in the absorption spectra. The NIR fluorescence decrease was also observed for **BC1** and **BC3** after aptamer binding. **BeCA** might partially quench the NIR fluorescence of these probes in the presence of Triton X-100. The decreased brightness of the RNA-fluorophore complex was probably due to the quenching interaction between the spatially adjacent guanine residues and the fluorophore[58, 109-110]. Nevertheless, **BC6** displayed an apparent decrease in the cyan channel and an increase in the NIR channel, showing the largest dynamic range (9.6-fold NIR/cyan ratio change) towards aptamer binding. **BC6** was therefore selected as the best color-shifting ligand for **BeCA** (Figure 44B).



	BC1	BC2	BC3	BC4	BC5	BC6	BC7
NIR turn-on	0.61	0.91	0.35	0.36	-	2.5	2.3
NIR/cyan change	-	3.2	-	2.4	-	9.6	4.2

*Figure 44. Fluorescence performance of **BC1-7** towards **BeCA** binding. (A) Emission spectra of **BC1-7** towards **BeCA** binding. **BC1-7** (1 μM) were incubated in the absence or presence of **BeCA** (25 μM) in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100. Excitations of 420 ± 5 nm and 630 ± 5 nm were used for the cyan and NIR windows, respectively. (B) The NIR turn-on factors were determined by dividing the maximal NIR fluorescence intensities of BC fluorophores in the presence of **BeCA** by those of BC fluorophores. NIR/cyan ratio change factors were determined by dividing the fluorescence ratios (maximal NIR fluorescence intensity/maximal cyan fluorescence intensity) of BC fluorophores in the presence of **BeCA** by those of BC fluorophores. The data was generated from panel A.*

3.1.6 Characterization of the color-shifting BeCA-BC6

BC6 outcompeted the other analogs and was selected as **BeCA**'s substrate for developing the color-shifting NIR FLAP. The spectroscopic and photophysical properties of **BeCA-BC6** were characterized.

3.1.6.1 Absorption and emission spectra

The absorption and emission spectra of **BC6** as a function of **BeCA** concentration were recorded. The absorbance of spirocyclic **BC6** at 420 nm decreased gradually while the absorbance of zwitterionic **BC6** at 665 nm rose as aptamer concentration increased, indicating that **BC6** favors the zwitterionic form over spirocyclic form in the presence of **BeCA** (Figure 45A). Consistent with the absorbance, the cyan fluorescence declined while NIR fluorescence boosted

when increasing the concentration of **BeCA** (Figure 45B). A 15-fold cyan/NIR fluorescence ratio (F_{480}/F_{684}) change was observed in the presence or absence of the **BeCA** (25 μM; Figure 45C).

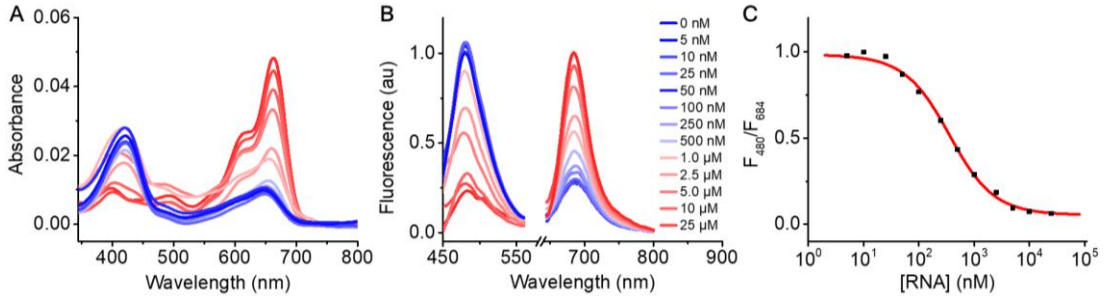


*Figure 45. Spectroscopic characterization of the **BeCA-BC6** complex. Absorption (A) and emission (B) spectra of **BeCA-BC6**. **BC6** (1 μM for absorption and 100 nM for emission spectra) was titrated with different amounts of **BeCA** (0 - 25 μM). (C) The plot of the emission ratio (cyan/NIR) as a function of **BeCA** concentration. The data points were generated using the fluorescence intensity ratios at the emission maxima of cyan and NIR windows, respectively (i.e., $F_{420\ nm}/F_{684\ nm}$). The spectra were measured in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100. Excitations of 420 ± 5 nm and 630 ± 5 nm were used for the cyan and NIR windows, respectively.*

3.1.6.2 Extinction coefficient and quantum yield

The extinction coefficients and quantum yields of **BC6** and the **BeCA-BC6** complex were further characterized. The absorption and emission spectra of **BC6** in buffer alone or under excess **BeCA** were measured with varying probe concentrations (Figure 46). **BC6** in HEPES buffer showed absorption peaks from both spirocyclic ($\lambda_{spirolactam\text{-}abs}$ = 420 nm) and zwitterionic ($\lambda_{zwitterion\text{-}abs}$ = 652 nm) forms. The absorption coefficients of **BC6** spirolactam and zwitterion were 27200 $mol^{-1}cm^{-1}$ and 11500 $mol^{-1}cm^{-1}$, respectively (Table 4). The quantum yields of **BC6** spirolactam and zwitterion were determined to be 0.04 and 0.08, respectively. In the presence of excess **BeCA**, **BC6** completely shifted to the zwitterionic form with a sharp absorption band, which bathochromic shifted to 665 nm. Meanwhile, negligible absorption of **BC6** spirolactam was detected. The absorption coefficient of **BC6** zwitterion in complex with **BeCA** was 52700 $mol^{-1}cm^{-1}$, 4.6-fold higher than that of **BC6** zwitterion in the buffer solution. The quantum yield of the **BeCA-BC6** complex in the NIR window was 0.07, similar to that of **BC6** in buffer alone. We concluded that the increased NIR brightness of **BC6** upon **BeCA** binding mostly was derived from the rise in molar extinction coefficient due to the equilibrium shift rather than the change in quantum yield.

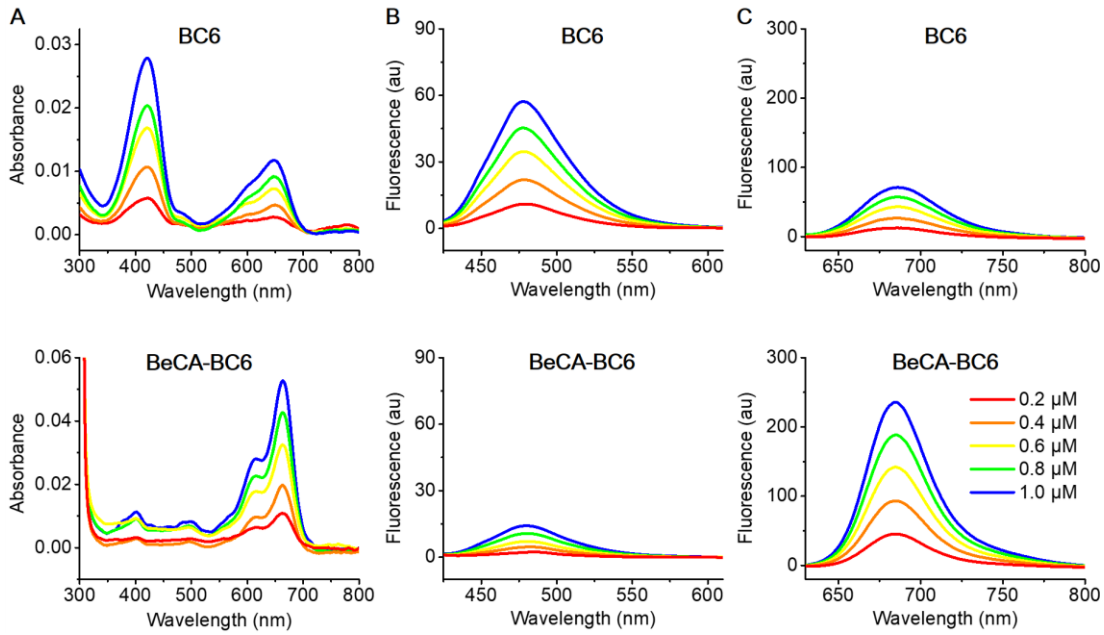


*Figure 46. Absorption (A), cyan emission (B), and NIR emission (C) spectra of **BC6** and **BeCA-BC6**. Different amounts of **BC6** (0 - 1 µM) were incubated without or with **BeCA** (25 µM) in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100. Excitations of 420 ± 5 nm and 615 ± 5 nm were used for the cyan and NIR windows, respectively.*

	λ_{abs} (nm)	λ_{em} (nm)	ε ($mol^{-1}cm^{-1}$)	Φ
BC6	420[a]/652[b]	478[a]/686[b]	27200[a]/11500[b]	0.04[a]/0.08[b]
BeCA-BC6	- [a]/665[b]	- [a]/684[b]	- [a]/52700[b]	- [a]/0.07[b]

*Table 4. Photophysical properties of **BC6** and **BeCA-BC6**. λ_{abs}: absorption maximum; λ_{em}: emission maximum; ε: molar extinction coefficient; Φ: quantum yield; [a] spirocyclic form; [b] zwitterionic form. The measurement was performed in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100.*

3.1.6.3 Dissociation constant

The binding affinity of **BeCA-BC6** was determined by plotting the NIR fluorescence at the emission maximum (684 nm) as a function of aptamer concentration. The fitted curve using the least-square equation revealed a K_D value of 220 nM (Figure 47A). Notably, when 0.05% Triton X-100 was present in the buffer system, a higher K_D value of 910 nM was observed. We assume that the interaction between the fluorophore and the detergent prevented the tight-binding of the fluorophore to **BeCA** (Figure 47B).

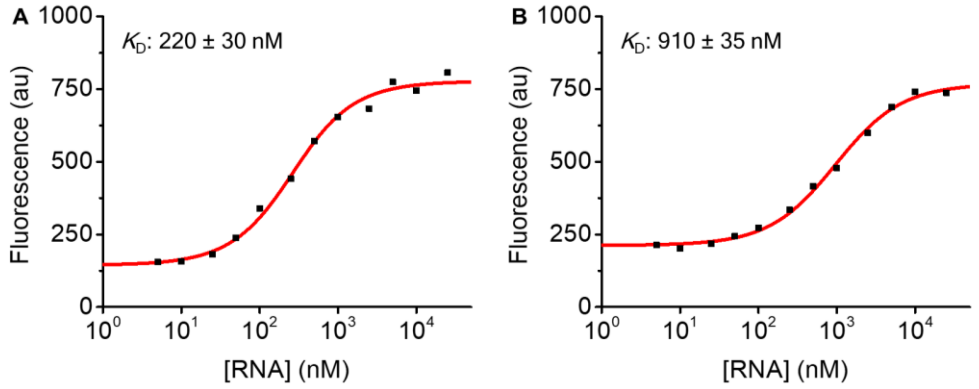


*Figure 47. K_D curves of **BeCA-BC6**. The K_D values were determined by titrating **BC6** (100 nM) with different amounts of **BeCA** (0 - 25 µM) in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl without (A) and with (B) 0.05% Triton X-100. Excitation wavelengths of 665 ± 5 nm and 630 ± 5 nm were used for (A) and (B), respectively.*

3.1.6.4 Temperature dependence

Temperature dependence study showed that **BeCA-BC6** exhibited a moderate temperature resistance *in vitro* as it retained 70% of the maximum fluorescence at the physiological temperature 37 °C, which is comparable to other reported thermal stable aptamer systems (Figure 48)[43, 58]. Interestingly, only 50% of the maximum fluorescence was retained for **BeCA-BC1-DN** at 37 °C. This indicated that the **BC6** could better stabilize the correct folding of **BeCA**, possibly through hydrogen bonding, electrostatic interactions, and Van der Waals forces compared to **BC1-DN**. This hysteresis phenomenon was also reported for other FLAPs such as Mango IV-TO1-B and Broccoli-BI[44, 48].

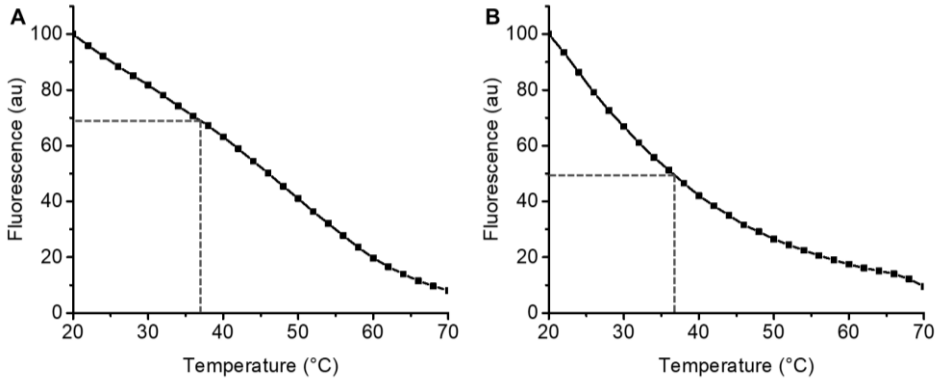


*Figure 48. Temperature dependence of **BeCA**. The temperature-dependent fluorescence of **BeCA-BC6** (A) and **BeCA-BC1-DN** (B). The fluorescence of the complexes (**BeCA**: 5 µM; fluorophores: 5 µM) were recorded in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl. The excitation wavelengths were 665 ± 5 nm and 652 ± 5 nm for **BeCA-BC6** and **BeCA-BC1-DN**, respectively. The recorded maximal fluorescence intensities were plotted against temperature.*

3.1.6.5 Cation dependence

Cations promote RNA folding by reducing the electrostatic repulsion between RNA phosphates. Among all biologically relevant cations, magnesium stabilizes the tertiary structures of most RNAs[118-119] while potassium ion promotes the formation of G-quadruplex[120-121]. Many FLAPs showed either magnesium or potassium dependent folding[45, 47, 58, 65]. Therefore, we examined the cation dependence of **BeCA-BC6**. The RNA-fluorophore complex retained 90% of the maximum fluorescence even in the absence of magnesium or potassium, indicating that the folding of **BeCA** was insensitive to these cations (Figure 49). This meant the folding of **BeCA** inside living cells would be less influenced by the intracellular concentration variations of the cations.

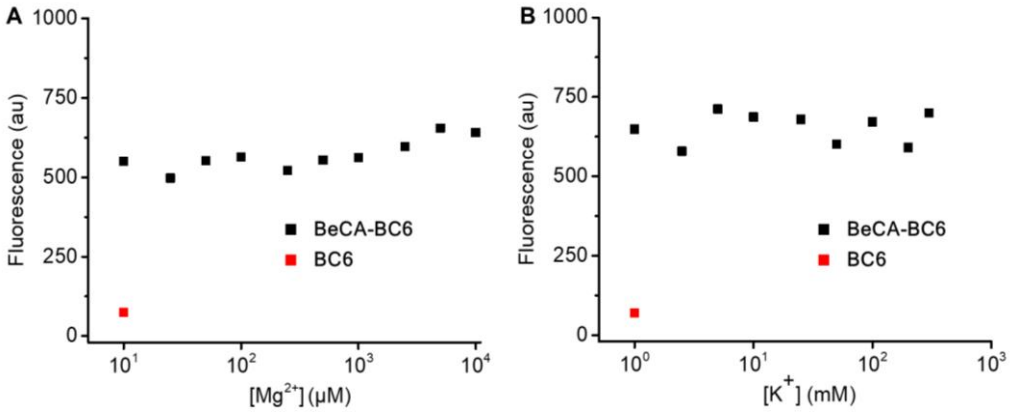


*Figure 49. Cation dependence of **BeCA-BC6**. (A) Magnesium dependence of **BeCA-BC6**. **BC6** (5 μM) was incubated with or without **BeCA** (5 μM) in 20 mM HEPES (pH 7.4) and 125 mM KCl containing varying concentrations of magnesium (0 – 10 mM). (B) Potassium dependence of **BeCA-BC6**. **BC6** (5 μM) was incubated with or without **BeCA** (5 μM) in 20 mM HEPES (pH 7.4) and 5 mM $MgCl_2$ containing varying concentrations of potassium (0 – 300 mM). For the fluorescence measurements, an excitation wavelength of 665 ± 5 nm was used, and the fluorescence intensities at the emission maximum (684 nm) were plotted against the magnesium/potassium concentration.*

3.1.7 Near-infrared imaging of mRNA in living bacteria

mRNAs generally possess short half-lives, complex structures and are pivotal in gene expression. Green fluorescent protein mRNA (*gfp*) was selected as a model RNA transcript to demonstrate the capability of **BeCA-BC6** as a NIR-fluorescent imaging tag in living cells.

3.1.7.1 Preparation of *pET-gfp-BeCA$_8$* plasmid

Eight copies of *BeCA* without any stabilizing scaffold was introduced into the 3'-untranslated region (UTR) of the gene coding GFP. The multimerized aptamer tag increases the signal-to-noise ratio when performing imaging[50, 65]. Four synonymous repeats of *BeCA* were chemically synthesized and cloned twice into the 1GFP vector. The synonymous design of the terminal helix of *BeCA* could reduce the probability of mispairing between the adjacent aptamer stems. The

sequencing result confirmed the successful insertion of *BeCA* into the *gfp* sequence.

pET-gfp-BeCA$_8$	5′ - …GTC GAC GGA GGT ACG TGG ACT GCT TAT GCA GTA GGA AGG CCT CC AAAA GGT CCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GGA CC AAAA GCT GCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GCA GC AAAA CGA GCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GCT CG CTC GAC AAG CTT GCG GCC GCA CTC GAC GGA GGT ACG TGG ACT GCT TAT GCA GTA GGA AGG CCT CC AAAA GGT CCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GGA CC AAAA GCT GCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GCA GC AAAA CGA GCT ACG TGG ACT GCT TAT GCA GTA GGA AGG GCT CG CTC GAG… - 3′

Table 5. The sequence of eight repeats of BeCA (indicated in red) in the pET-gfp-BeCA$_8$ plasmid.

3.1.7.2 Near-infrared imaging of *gfp-BeCA$_8$* mRNA

The engineered *pET-gfp-BeCA$_8$* plasmid and the *pET-gfp* (1GFP) plasmid as control were transformed into *E. coli* cells. The transcription of *gfp-BeCA$_8$* and *gfp* mRNA was under the control of a lac operon, which could be effectively activated by isopropyl-β-d-thiogalactopyranoside (IPTG). The mRNAs were expressed with IPTG induction and imaged in the presence of **BC6** in *E. coli* at room temperature (Figure 50). Similar expression levels of GFP were detected in both groups with and without *BeCA* tags, suggesting no obvious influence of the aptamer tag on the mRNA translational efficiency. In the meantime, bright NIR fluorescence predominantly at the bacterial poles was detected in bacterial cells expressing *gfp-BeCA$_8$* but not in the control cells where only *gfp* mRNA was present. This indicated *BeCA* could selectively bind **BC6** zwitterion and activates the NIR fluorescence. The accumulation of mRNA at bacterial poles demonstrated by the NIR signal was likely due to the localization of pET plasmid where mRNA transcription initiates[58, 122]. The average NIR fluorescence intensities of the bacterial poles of *gfp-BeCA$_8$*-expressing cells were 6.4-fold, 6.3-fold, and 5.3-fold higher than those of the control bacteria at the probe concentrations of 200, 500, and 1000 nM, respectively. The decreased signal-to-background ratio at high probe concentration (1 μM) was presumably due to the elevated background noise from the unbound **BC6**.

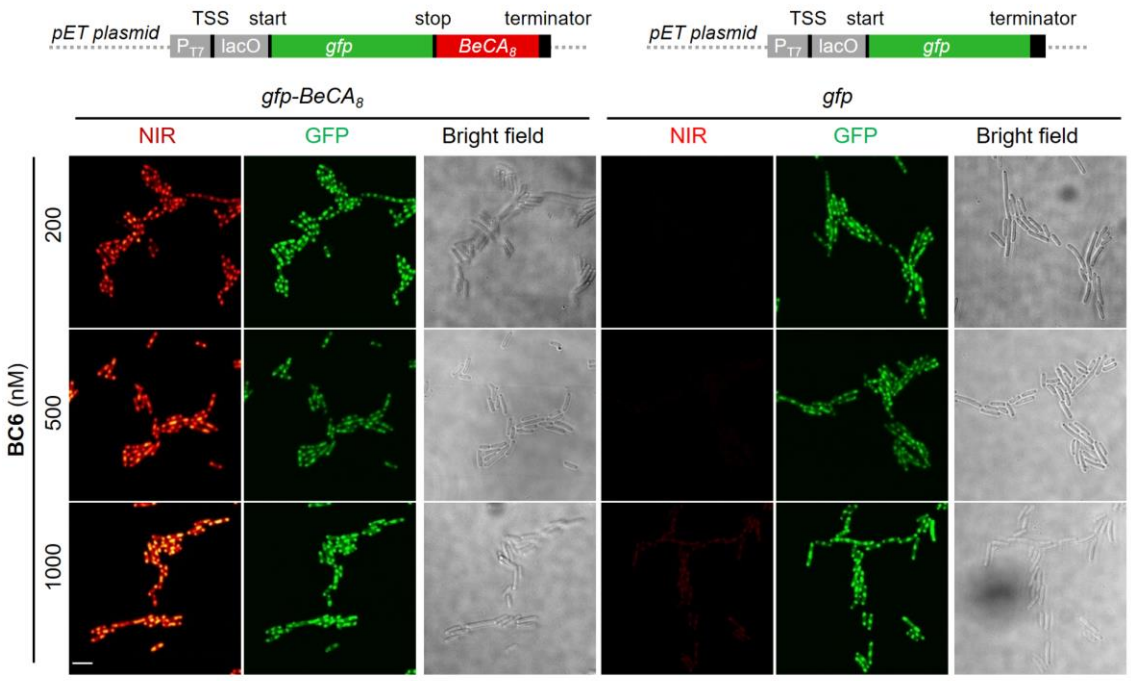


Figure 50. Live-cell confocal images of gfp mRNA tagged with BeCA in E. coli. Bacteria were transformed with either pET-gfp-BeCA$_8$ or pET-gfp plasmid. The mRNA expression was induced by the addition of IPTG (1 mM). Cells were incubated with ***BC6*** *(200, 500, 1000 nM) for 15 min and imaged. Green and NIR channel images were acquired using 488 nm and 640 nm lasers, respectively. Scale bar: 5 µm.*

3.1.8 Ratiometric imaging of tRNA in living bacteria

One unique feature of **BeCA-BC6** is the dual-color nature of the system. Proof-of-principle live-cell imaging experiments in *E. coli* were performed to test whether the **BeCA-BC6** system could be used as a color-shifting NIR FLAP for ratiometric imaging and analysis of target RNAs.

3.1.8.1 Preparation of *pET-tRNA-BeCA* plasmid

A tRNA$_{Lys}$ tagged with *BeCA* was selected as a model RNA transcript for the ratiometric RNA imaging in living cells. We inserted the *BeCA* sequence to the anticodon region of the tRNA$_{Lys}$. Thus, the *BeCA* tag enabled the specific labeling of the recombinant tRNA while the tRNA scaffold promoted the correct folding of *BeCA* in cells due to its structural stability[56, 65, 123]. The *tRNA-BeCA* was prepared and cloned into a pET-28a(+) vector.

pET-tRNA-BeCA	5′ - …TAA TAC GAC TCA CTA TA GCG GAT TTA GCT CAG TTG GGA GAG CGC CAG GGA GGT ACG TGG ACT GCT TAT GCA GTA GGA AGG CCT CC CTG GAG GTC CTG TGT TCG ATC CAC AGA ATT CGC ACC A… - 3′

Table 6. The sequence of tRNA-BeCA in the pET-tRNA-BeCA plasmid. T7 promoter, tRNA scaffold, BeCA sequences are indicated in blue, green, and red, respectively

3.1.8.2 Imaging of *tRNA-BeCA* transcript

The engineered *pET-tRNA-BeCA* plasmid and the *pET-tRNA*[d] plasmid as control were transformed into *E. coli* and expressed with IPTG induction. The cells expressing *tRNA-BeCA* or *tRNA* were incubated with varying concentrations of **BC6** and imaged in the NIR and cyan channels simultaneously at 37 °C (Figure 51). Interestingly, we observed similar cyan signals in both groups with or without *BeCA*, and the intracellular cyan signal increased as probe concentration increased. This was attributed to the dynamic equilibrium between the intracellular and extracellular spirocyclic **BC6**, which was assumed to possess higher cell permeability than its zwitterionic counterpart[82]. On the contrary, the *tRNA-BeCA*-expressing cells showed significantly higher NIR signals than the control cells at each probe concentration condition due to the specific interaction between *BeCA* and **BC6** zwitterion. Increasing the probe concentration from 200 to 500 nM, as expected, resulted in a rise in the NIR signal. However, a further increment of the probe concentration to 1 μM showed negligible NIR fluorescence improvement, yet with elevated background signal as evidenced by the signal in the control cells. This implied that the binding of **BC6** to *tRNA-BeCA* was already saturated at 500 nM dye concentration.

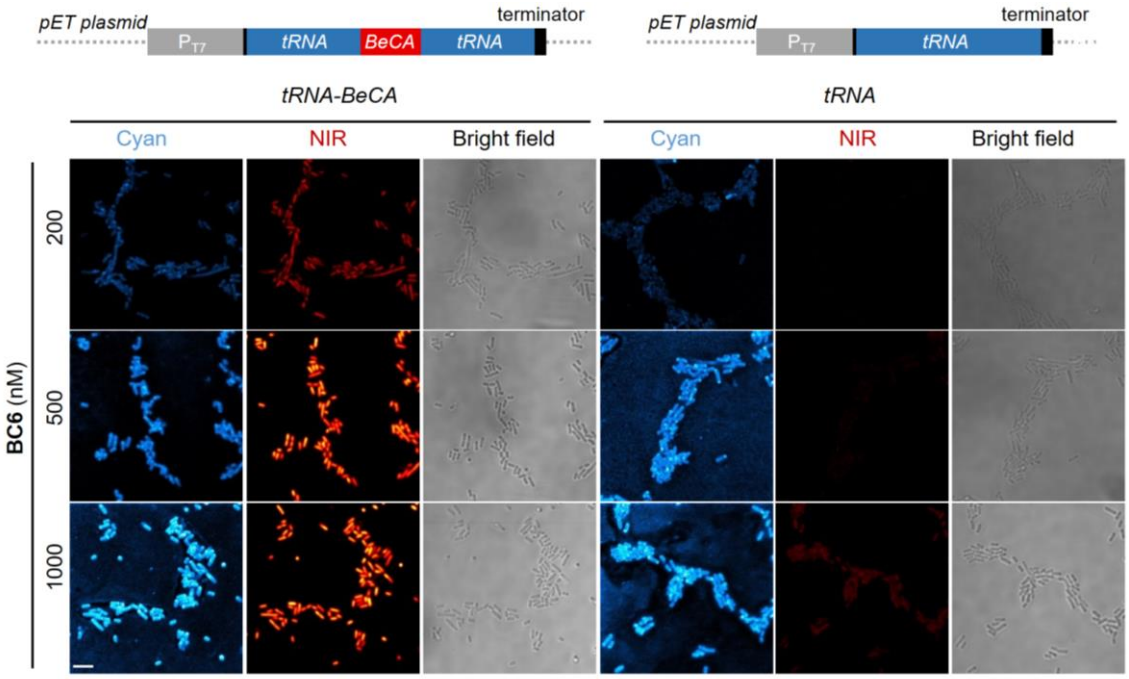


Figure 51. Live-cell confocal images of BeCA-tagged tRNA in E. coli. pET-tRNA-BeCA plasmid and pET-tRNA plasmid as control were transformed into bacteria. The RNA expression was induced by the addition of IPTG (1 mM). The cells were incubated with ***BC6*** *(200 - 1000 nM) and imaged. Cyan and NIR channel images were acquired using 405 nm and 640 nm lasers, respectively. Scale bar: 5 μm.*

[d] This *pET-tRNA* (or *pET-tRNA*$_{Lys}$) plasmid was cloned and kindly provided by Dr. Murat Sunbul.

3.1.8.3 Ratiometric imaging of tRNA at different expression levels

Next, we examined the capability of the color-shifting **BeCA-BC6** for analyzing the abundance of a target RNA (i.e., *tRNA-BeCA*) in living cells. The expression level of *tRNA-BeCA* was controlled by the amount of IPTG added, and the *tRNA-BeCA*-expressing cells were imaged in both cyan and NIR channels simultaneously (Figure 52). The ratiometric images were obtained by pixel-to-pixel intensity division of NIR to cyan channels.

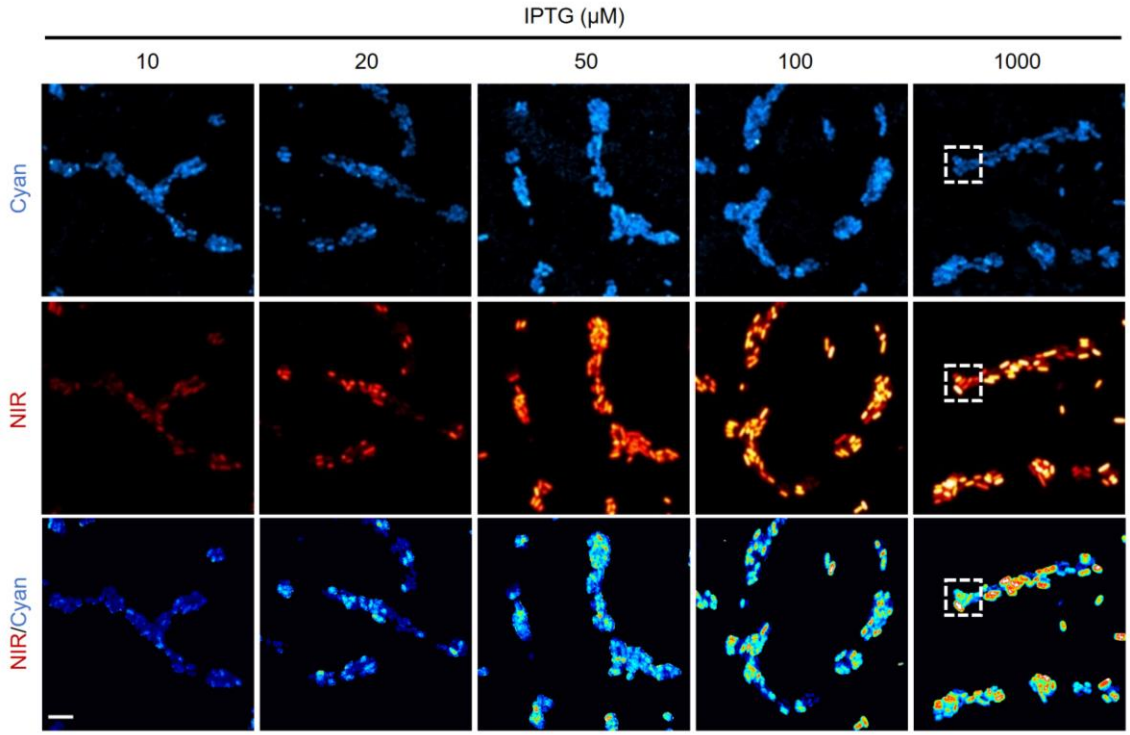


Figure 52. Live-cell confocal images of tRNA-BeCA expressed at different levels in E. coli. IPTG induced the transcription of tRNA-BeCA in E. coli transformants with varying concentrations (10 – 1000 μM). The cells were imaged after incubating with ***BC6*** *(200 nM) for 15 min. Cyan and NIR channel images were acquired using 405 nm and 640 nm lasers, respectively. The ratiometric images (NIR/cyan) were obtained from the pixel-to-pixel intensity ratio values from the NIR and cyan channel images. Scale bar: 5 μm.*

The NIR to cyan intensity ratios (F_{NIR}/F_{cyan}) of the representative cells quantified from the ratiometric images of each group were plotted (Figure 53A). The mean ratio values were positively correlated with IPTG concentrations, showing a 1.7-, 2.4-, 3.0- and 3.7-fold increase at IPTG concentrations of 20, 50, 100, and 1000 μM, respectively, relative to that at IPTG concentration of 10 μM. To evaluate whether these values obtained from the ratiometric images could represent the intracellular abundance of *tRNA-BeCA*, we isolated and quantified *tRNA-BeCA in vitro* from bacteria with the same treatments. The *tRNA-BeCA* transcript was extracted from the lysed cells and quantified by gel electrophoresis (Figure 53B). The band volume intensities of *tRNA-BeCA* from bacteria groups treated with 20, 50, 100, 1000 μM of IPTG were determined to be 2.0-, 2.9-, 3.3-, and 3.7-fold higher

than that from bacteria group treated with 10 μM of IPTG (Figure 53C). The *in vitro* data matched well with the live-cell imaging data, demonstrating the utility of **BeCA-BC6** to predict the expressing level of target RNA in living cells.

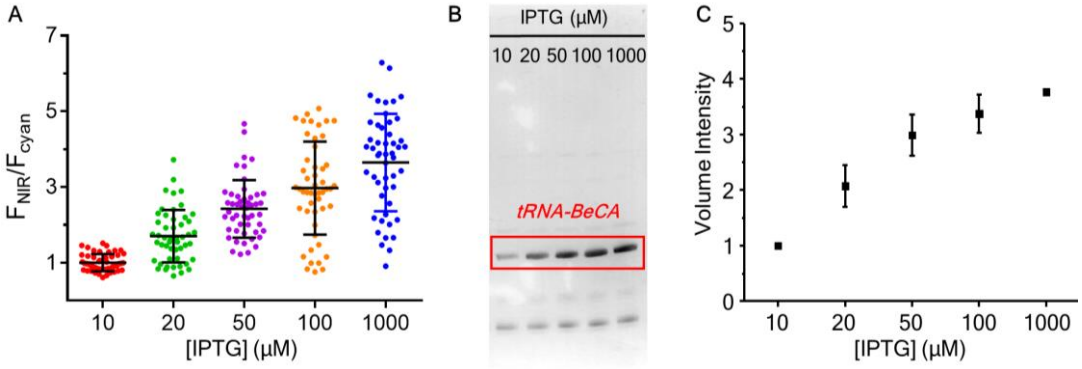


Figure 53. Quantification of intracellular tRNA-BeCA amount in E. coli. (A) Quantification of the mean fluorescence ratio (F_{NIR}/F_{cyan}) of bacteria expressing tRNA-BeCA at different IPTG concentrations (10 – 1000 μM) from the ratiometric images. Each dot represents a single cell. Means ± standard deviation (s.d.) are shown (n = 50 cells from five independent images). The mean of the F_{NIR}/F_{cyan} ratio for bacteria treated with 10 μM was normalized to 1. (B) Denaturing polyacrylamide gel analysis of BeCA in total RNA isolated from bacteria treated with different IPTG concentrations (10 – 1000 μM). Under each condition, the same number of bacteria were used to isolate total RNA. OD_{600} values were used to quantify the number of bacteria. (C) Normalized band volume intensity of tRNA-BeCA. The band intensity of each sample was normalized to that of the sample treated with 10 μM of IPTG. Mean ± standard deviation (s.d.) are shown (n = 2 gels from two independent experiments).

Besides a quick overview of intracellular RNA abundance at the population (bulk) level, the ratiometric images could provide more precise information at the single-cell level. For example, cell a and cell b in Figure 54 exhibited different NIR fluorescence intensities, while the signal in the cyan channel indicated that this NIR signal fluctuation might be derived from variations in intracellular probe amount or focus layer. The ratiometric signals indicated that they could have similar levels of intracellular *BeCA*-tagged tRNA, thereby offering correction for this discrepancy. Extensive heterogeneity of *tRNA-BeCA* transcription at the single-cell level in *E. coli* due to stochasticity in transcriptional regulation could also be revealed using the ratiometric images, especially at high IPTG concentration[124]. For example, cell c showed a high NIR/cyan intensity ratio while cell d and cell e only held medium and low ratios, indicating the different *tRNA-BeCA* transcription efficiency of individual cells under the same treatment (Figure 54).

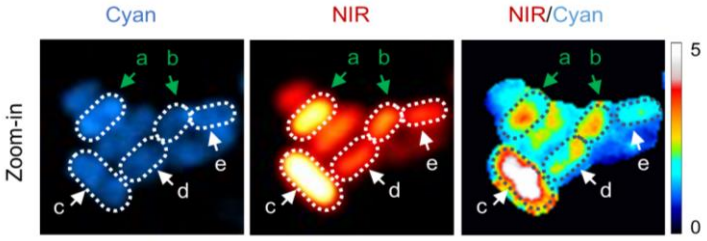


Figure 54. Zoom-in images of the tRNA-BeCA-expressing cells depicted in Figure 52. Green arrows indicate the cells a and b, of which intracellular tRNA-BeCA concentration was corrected using ratiometric analysis. White arrows show transcription heterogeneity at the single-cell level. Transcription efficiency according to the ratiometric images: cell c > cell d > cell e.

3.1.9 Development of a ratiometric sensor platform based on the color-shifting BeCA-BC6

Similar to fluorescent protein-based protein sensors that restore/change fluorescence emission upon analytes sensing, RNA-based sensors using FLAPs as fluorescence reporters have been developed for various small molecules[125-127]. Ratiometric sensing possesses unique advantages in accurate quantification compared to standard fluorescent light-up sensing (chapter 2.1). Therefore, we developed a ratiometric sensor platform by utilizing the color-shifting module **BeCA-BC6**.

3.1.9.1 Design of a tetracycline sensor

Using tetracycline (Tc) as a model analyte, a Tc sensor was designed (Figure 55). The sensor RNA comprises a recognition module, a transducer module, and a fluorescence reporter module. The Tc-binding aptamer as the recognition module would be fused to the fluorescence reporter **BeCA** via a transducer stem. This sensor RNA would be incorrectly folded in the absence of target (i.e., Tc), showing cyan fluorescence only from the unbound **BC6**. However, in the presence of Tc, the binding of Tc to its cognate aptamer would promote the stabilization of the transducer stem and induce the correct folding of **BeCA**. The refolded **BeCA** aptamer would form a NIR-fluorescent complex with **BC6**. The ratio of the two emissions could be used to determine the Tc concentration.

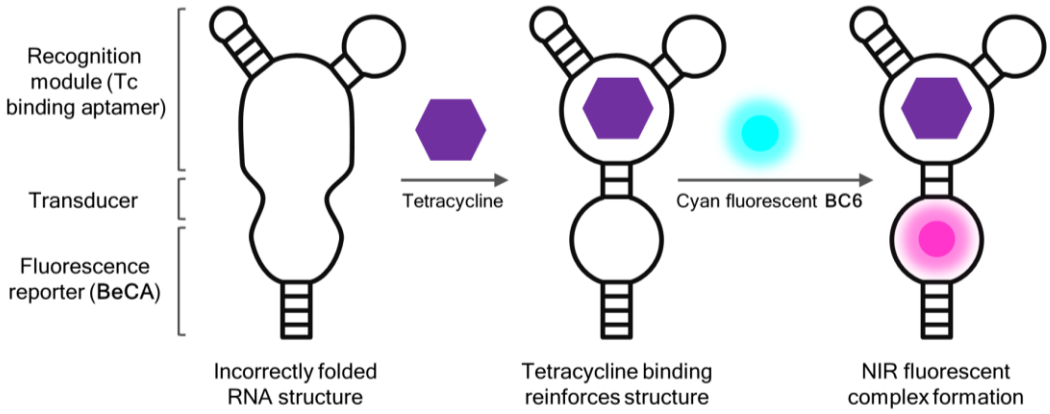


*Figure 55. Design of a ratiometric sensor based on **BeCA-BC6**. The sensor RNA comprises a recognition module, a transducer module, and a fluorescence reporter module. Tetracycline binding to its recognition aptamer reinforces the correct folding of **BeCA** and shifts the fluorescence of **BC6** from cyan to NIR.*

3.1.9.2 Identification of a fusion point

For insertion of the recognition module (i.e., Tc-binding aptamer) to **BeCA**, an entry point needs to be determined. The structural study of **BeCA** (chapter 3.1.3.5) indicated that the main loop was the binding core while the terminal stem-loop had minimal effect on the complexation between **BeCA** and **BC6**. Based on these observations, two **BeCA** variants were generated by replacing the internal stem-loop with either a replacement or a mismatched stem-loop (Figure 56A). The replacement stem-loop possesses a different nucleotide sequence but forms the same structure. Variant bearing the replacement loop showed a similar NIR turn-on upon **BC6** binding compared to the original **BeCA** (Figure 56B). Nevertheless, the introduction of the mismatched stem impeded the correct folding of **BeCA** due to the reduced stem stability and consequently led to a reduced NIR signal in the presence of **BC6**. Therefore, we concluded that the internal helix is essential for **BeCA** in a non-sequence-specific manner, making it a good fusion point for incorporating a Tc-binding aptamer.

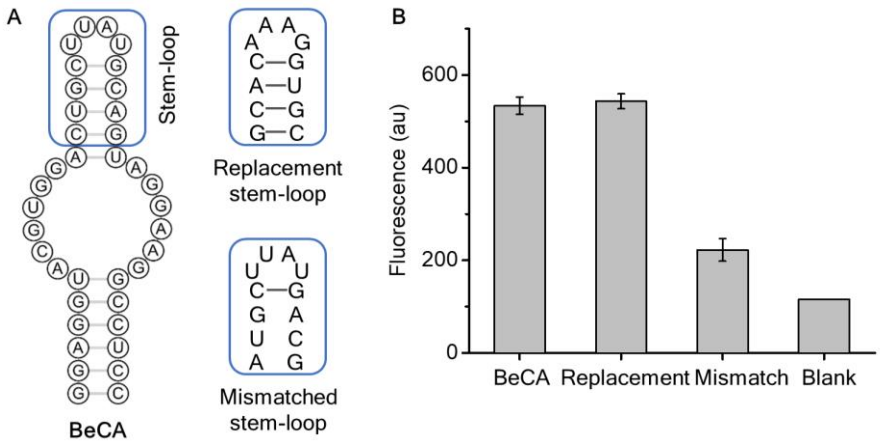


*Figure 56. Identification of a fusion point of **BeCA**. (A) The predicted secondary structure of **BeCA**. The internal, replacement, and mismatched stem-loops were indicated with blue boxes. (B) Fluorescence intensities of original, replacement, and mismatched **BeCA**. **BC6** (100 nM) was incubated with 5 μM of original, replacement, mismatched **BeCA**s or in buffer alone, and the maximal fluorescence intensities were recorded with an excitation of 665 ± 5 nm in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl.*

3.1.9.3 Transducer design and screening

We selected a Tc-binding RNA aptamer with a binding affinity in the picomolar range and inserted it into **BeCA** via the internal stem (Figure 57A)[128]. The shared stem was redesigned to function as a thermodynamically unstable transducer module. Two approaches were applied to destabilize the transducer stem, and a total of eight variants containing different transducer sequences were generated (Figure 57B). For transducer 1 to 4, mismatched base pairs such as A-A or weak base pairs such as G-U were introduced. For transducer 5 to 8, short stems ranging from 1 to 3 base pairs were used instead of 4 base pairs in the original stem.

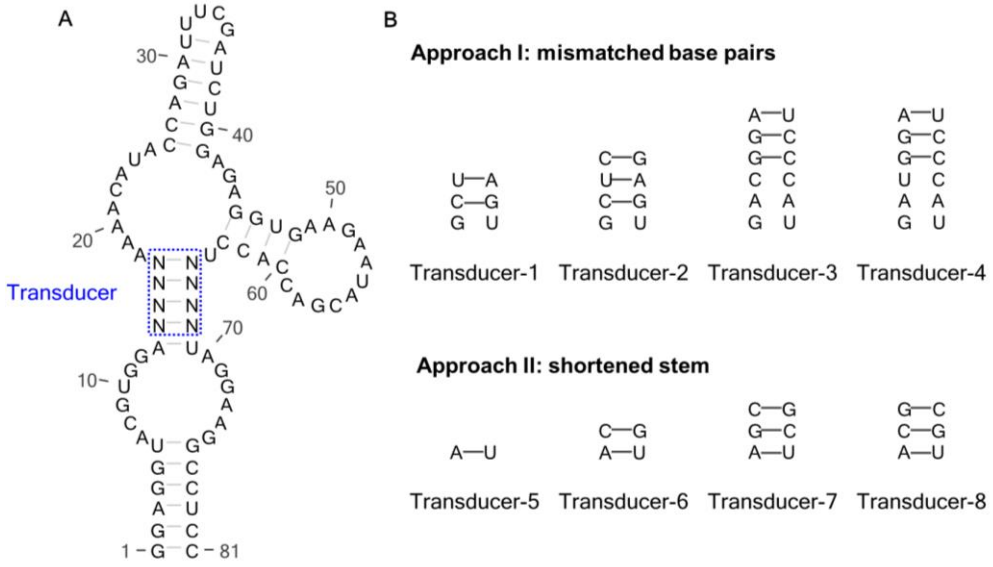


*Figure 57. Design of the **BeCA**-based Tc sensor. (A) The sequence and predicted secondary structure of the Tc sensor. The Tc-binding aptamer and **BeCA** were fused via a transducer stem (boxed in blue). N represents different possible nucleotides. (B) The sequences of eight transducer stems based on two design strategies.*

The eight RNA sensors were assayed for their fluorescence sensing properties. In the absence of Tc, RNAs consisting of transducer 3 (**T3**), transducer 4 (**T4**), and transducer 5 (**T5**) exhibited significant lower initial NIR fluorescence compared to the other RNA variants (**T1**, **T2**, **T6**, **T7**, and **T8**), indicating that the transducer design of **T3**, **T4** and **T5** disturbed the folding of **BeCA** at a higher extent (Figure 58). The subsequent addition of Tc led to 1.6- and 1.9-fold fluorescence increases for **T3** and **T5**, respectively, while the rest displayed minor fluorescence changes. The NIR signal rise of **T3** and **T5** was presumably due to the restoration of **BeCA** induced by the Tc binding to its cognate aptamer. However, the turn-on factors were not as high as the theoretical maximum (The NIR turn-on factor of **BC6** upon **BeCA** binding) owing to the incomplete disruption and/or reconstruction of **BeCA** in the absence or presence of Tc. All in all, **T5** displayed the maximal NIR fluorescence turn-on among all the designed sensor variants upon Tc sensing and was therefore chosen for further spectroscopic characterization.

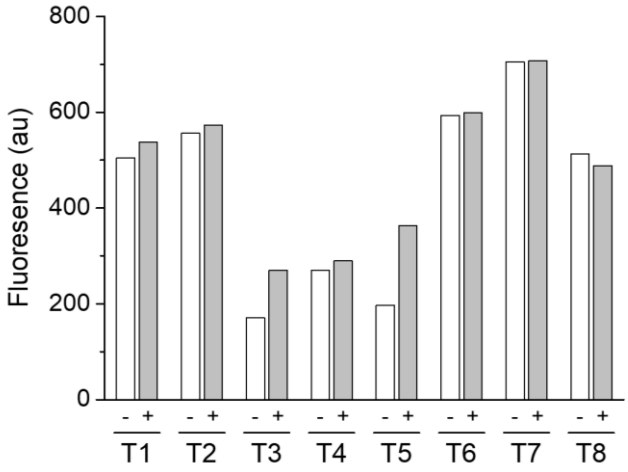


*Figure 58. Fluorescence screening of Tc sensors **T1-8**. **T1-8** (5 µM) were individually incubated without (-) or with (+) Tc (200 µM) in the presence of **BC6** (100 nM). The maximal fluorescence intensities were recorded with an excitation wavelength of 665 ± 5 nm in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, and 125 mM KCl.*

3.1.9.4 Spectroscopic characterization of T5-BC6

Spectroscopic characterization of **T5-BC6** as a ratiometric sensor for Tc was executed. **T5-BC6** was incubated with varying concentrations of Tc, and the emissions in both cyan and NIR windows were recorded (Figure 59A,B). The fluorescence ratio of the two emissions was plotted against the Tc concentration. The fitted curve revealed a half-maximal effective concentration (EC_{50}) of 470 nM, a dynamic range of 2.3-fold, and a detection range of 0.2 - 1 µM (determined as 10% to 90% of the ratio output; Figure 59C).

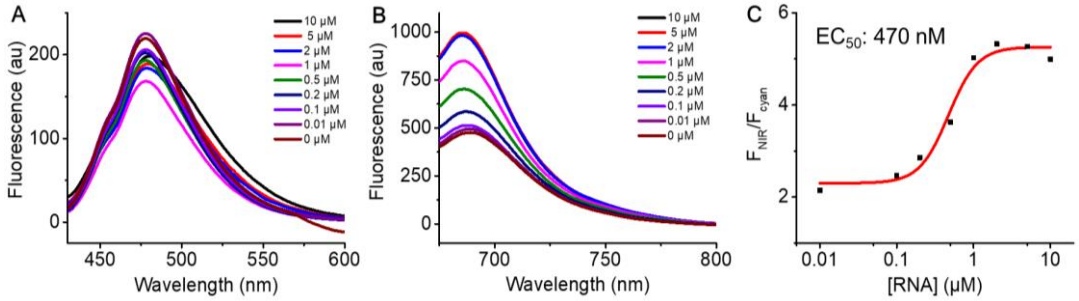


*Figure 59. Characterization of **T5-BC6**. Cyan (A) and NIR (B) emission spectra of **T5-BC6** (**T5**: 5 µM; **BC6**: 500 nM) incubated with different amounts of Tc (0 - 10 µM). (C) The plot of the emission ratio (F_{NIR}/F_{cyan}) as a function of **T5** concentration. The data points were generated using the ratios of the maximal cyan fluorescence intensities to maximal NIR fluorescence intensities at each **T5** concentration condition. The spectra were measured in 1 × ASB containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% Triton X-100. Excitations of 420 ± 5 nm and 630 ± 5 nm were used for cyan and NIR windows, respectively.*

Although the ratiometric Tc sensor (**T5-BC6**) was proven to work *in vitro*, the low dynamic range (2.3-fold) limited its usage in living cells. The initial trial of **T5-BC6** in *E. coli* cells demonstrated only a very slight NIR fluorescence increase (1.2-fold) under high Tc concentration (10 µM; data not shown). Further optimization

of this ratiometric sensor that allows a higher brightness and a larger dynamic range is required to achieve a better sensing functionality in cells.
In conclusion, in this chapter (chapter 3.1), we successfully developed a color-shifting aptamer-fluorophore module **BeCA-BC6** and demonstrated its capability in RNA imaging and analyte sensing. This new platform showed the first example of imaging both the unbound fluorophore and the aptamer-fluorophore complex simultaneously, opening up new possibilities for quantitative monitoring of target RNAs in living cells using the ratiometric signal.

3.2 Evolution of silicon rhodamine-binding near-infrared fluorescent light-up aptamer through fluorescence-activated cell sorting

Spirocyclization-based fluorogenic dyes (e.g., SiR, JF, and MaP) are interesting substrates for developing new FLAPs due to their excellent cell permeability and photophysical properties[73, 81-82]. The *in vitro* selection method SELEX is unfortunately not ideal for generating FLAPs for this kind of fluorophore. The affinity-oriented selection is would potentially lead to a loss of bright aptamers during the course of selection. FACS is a good alternative as both affinity and brightness of a FLAP could be assessed simultaneously during the selection.
SiRA is a reported NIR FLAP using the spirocyclization-based SiR as the fluorogenic ligand. Unfortunately, it is currently incapable of imaging RNA in mammalian cells, presumably due to the low binding affinity (K_D = 430 nM)[65]. Taking advantage of FACS, we evolved new SiR-binding RNA sequences, which displayed improved fluorescence performances compared to SiRA. All the results will be described and discussed in detail in the following texts.

3.2.1 Method development and validation

FACS is a valuable tool for detecting and separating cell populations of interest based on their specific physical and optical properties. Many antibodies have been evolved through FACS-based screening in yeast and mammalian cells[129-130]. Nevertheless, its application to bacteria is still challenging due to the difficulty in discriminating small bacteria from noise[131-132]. We first established a standard protocol for bacteria sorting and validated its utility for isolating fluorescent aptamer-expressing bacteria from a mixed population.

3.2.1.1 Operation principles of FACS-based bacteria sorting

The key to successful bacteria sorting using FACS includes two parts. (1) Instrument setup. The most obvious and prominent difference of bacteria is the cell diameter (< 1 µm) compared to eukaryotic cells (10 – 100 µm). Therefore, the thresholds and PMT voltages on the forward scatter (FSC) and side scatter (SSC) should be correctly set for identifying bacteria and are distinct from those used for detecting mammalian cells. Besides, the small size of bacteria necessitates the viewing of FSC-SSC scatter plot in dual-logarithmic scales. Otherwise, the bacteria populations of interest would be buried in the corner. (2) Sample preparation. The signal of bacteria in the scatter plot is close to the electronic noise of the machine. That is why the identification of the bacteria population is difficult. Using a fluorescent marker or an appropriate control could help define the microbial population. To reduce background noise, filtering the buffers through a 0.2 µM filter and using buffers with a similar refractive index as "sheath" are recommended. A cell density between 5×10^5 to 10^7 cells/mL and a slow flow rate around 3000 events/sec of the sample are preferred to avoid cell

aggregation. All the following experiments were performed based on the proper settings according to these principles.

3.2.1.2 Mock GFP-expressing bacteria sorting

Mock GFP-expressing bacteria sorting was carried out to examine the feasibility of FACS-based bacteria sorting. The commercial 1GFP plasmid and a tRNA expression plasmid (*pET-tRNA*) as negative control were transformed into BL21 *E. coli* cells. GFP positive (+) and negative (-) cells were mixed in a 1 : 10 ratio and subjected to the cell sorter. Nice separation between the two populations was observed on the analyzer (Figure 60). The percentage of GFP+ cells was 4.3%, slightly lower than the theoretical number (9%) due to the counted "fake" negative events from the noise. The two subsets, GFP+ and GFP- cells, were collected, respectively, and subjected to reanalysis immediately. An increase from 4.29% to 72.3% of Q2 was observed for the sorted GFP+ cells, while the number rose from 95.7% to 100% of Q3 for the sorted GFP- cells. Successful enrichment of populations of interest was achieved by gated sorting (Figure 60).

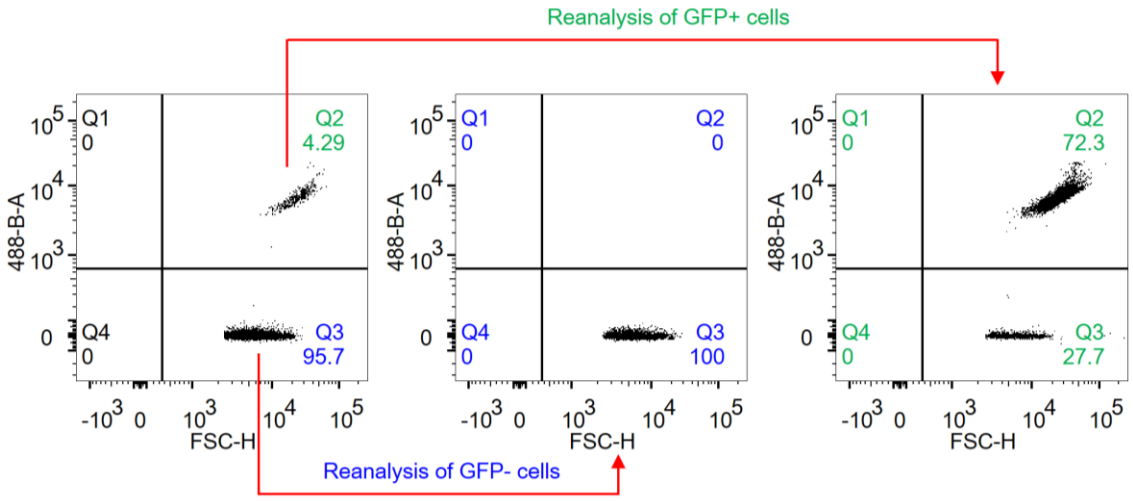


Figure 60. Flow cytometry analysis of GFP-expressing bacteria. 1GFP plasmid and pET-tRNA plasmid as control were transformed into E. coli cells. The cells were treated with 1 mM of ITPG and mixed in the ratio 1 : 10 (GFP+ : GFP-). The sorted GFP+ and GFP- cells were subjected to reanalysis. The cells were excited with a 488 nm laser, and the scatter plots of FSC vs. fluorescence are shown. Each dot represents one event, and a total of 10^4 events are displayed. The same gate was applied for the three plots, and the fractions of each quadrant are indicated.

A distinct separation between "positive" and "negative" cells is not always easily achieved in the real sorting. An ambiguous boundary or even overlapping between populations is often seen at the initial rounds of FACS-based selection. It is challenging to set an appropriate sorting gate under this condition. To mimic such an issue, we regulated the GFP expression level in *E. coli* cells mixed the GFP+ and GFP- cells at a ratio of 1 : 100. As expected, the partitioning of two subgroups was easily achieved for cells treated with high IPTG concentrations (1 and 0.1 mM), where high expression of GFP was induced. However, it became

extremely difficult for cells treated with low (0.01 mM) and no IPTG where low expression of GFP was present (Figure 61A). The GFP+ cells overlapped with GFP- cells to certain extents in the scatter plots. Nevertheless, we set a gate where the majority of the cells were present in the Q3 and isolated the brightest cells in Q2 for each IPTG-treated group. The collected cells were placed on the selective agar plates and analyzed by green/white colony counting. As shown in Figure 61B, the percentages of GFP+ cells for all four groups significantly increased after sorting. This demonstrated that a satisfying recovery efficiency of a subpopulation could be achieved by proper gating.

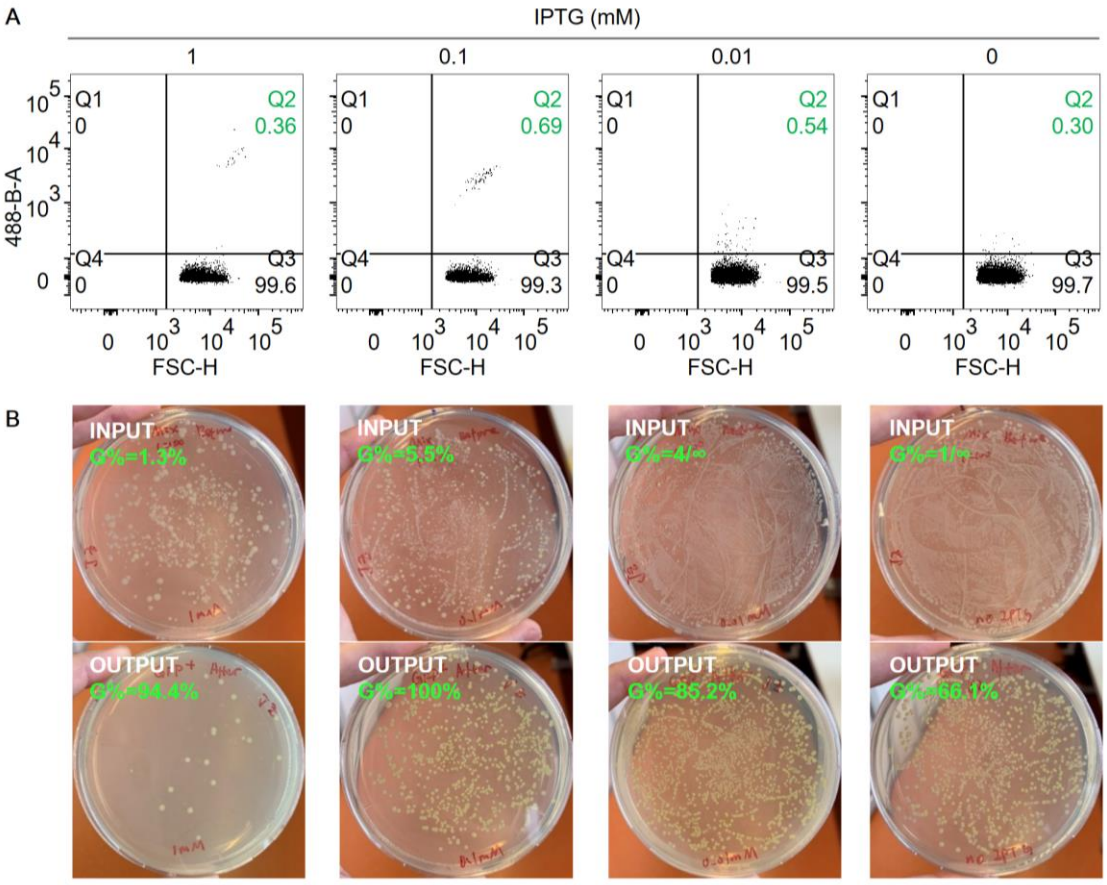


Figure 61. The sensitivity of FACS for bacteria sorting. (A) Flow cytometry analysis of GFP-expressing bacteria. 1GFP plasmid and pET-tRNA plasmid as control were transformed into E. coli cells. The cells were treated with different amounts of ITPG (0 - 1 mM), mixed in the ratio 1 : 100 (GFP+ : GFP-), and subjected to the cell sorter. The cells were excited with a 488 nm laser, and the scatter plots of FSC vs. fluorescence are shown. Each dot represents one event, and a total of 10^4 events are displayed. The fractions of each quadrant are indicated. The same sorting gate was applied to each IPTG-treated group, and the highly fluorescent cells in Q2 were collected. (B) Green/white colony screening result. The same quantities of the input bacteria mixtures (~10^6 cells) and the sorted highly fluorescent cells (5000 events) were placed on LB-kanamycin agar plates. Green and white colonies represent GFP+ and GFP- cells, respectively. The percentages of green colonies over total colonies (white + green) were indicated in green.

It is worth noting that high IPTG concentration could potentially cause toxicity to BL21 *E. coli* cells[133]. As shown in Figure 61B, the groups treated with low IPTG concentrations had significantly more colonies than the groups treated with high

IPTG concentrations, even though we placed the same amount of bacteria on the plates. This IPTG related cytotoxicity was further investigated in the following chapter.

3.2.1.3 Assessment of bacterial viability

A better signal separation between subpopulations is always preferred for FACS-based screening of aptamers, especially at the starting rounds. Highly over-expressed aptamers in cells could saturate the fluorogenic ligands and initiate the maximal fluorescence enhancement, thereby allowing the best separation between "positive" and "negative" cells. BL21 Star (DE3) *E. coli* strain has an increased RNA stability owing to the mutated *rne* gene, which reduces the number of endogenous ribonucleases[134]. pET vectors (plasmid for expression by T7 RNA polymerase) carrying the genes of interest driven by T7 promoter are recognized as the best vector series for bacterial protein/RNA expression due to the high activity and selectivity of phage T7 polymerase[135]. Considering the extremely high expressing level, pET expression vectors and BL21 Star (DE3) as host strains are preferable for in-bacteria aptamer selection.
Keeping cells in their healthy living status is pivotal in classic FACS since the cells need to be recycled for growth or physiological analysis after sorting. Unfortunately, cellular toxicity was observed for pET-transformed BL21 cells at high IPTG concentrations (chapter 3.2.1.2). This could impede the real FACS-based evolution of aptamer in living bacteria since the collected cells under the toxic status would die rather than grow and replicate. To investigate the toxicity issue, we assessed the viability of cells transformed with different expression vectors. "Empty" BL21 cells showed a high survival rate (> 80%) for all IPTG-treated groups, indicating no apparent toxicity of ITPG itself to bacterial cells (Figure 62A). However, BL21 cells transformed with either *pET-GFP* or *pET-RhoBAST* [e] displayed a significant drop in cell viability when the IPTG concentration increased to 40 µM or above (Figure 62B,C). This demonstrated BL21 cells transformed with pET expression vectors suffered from cytotoxicity, which might be derived from the metabolic burden induced by high amounts of IPTG[133].

[e] This plasmid was cloned and kindly provided by Dr. Murat Sunbul.

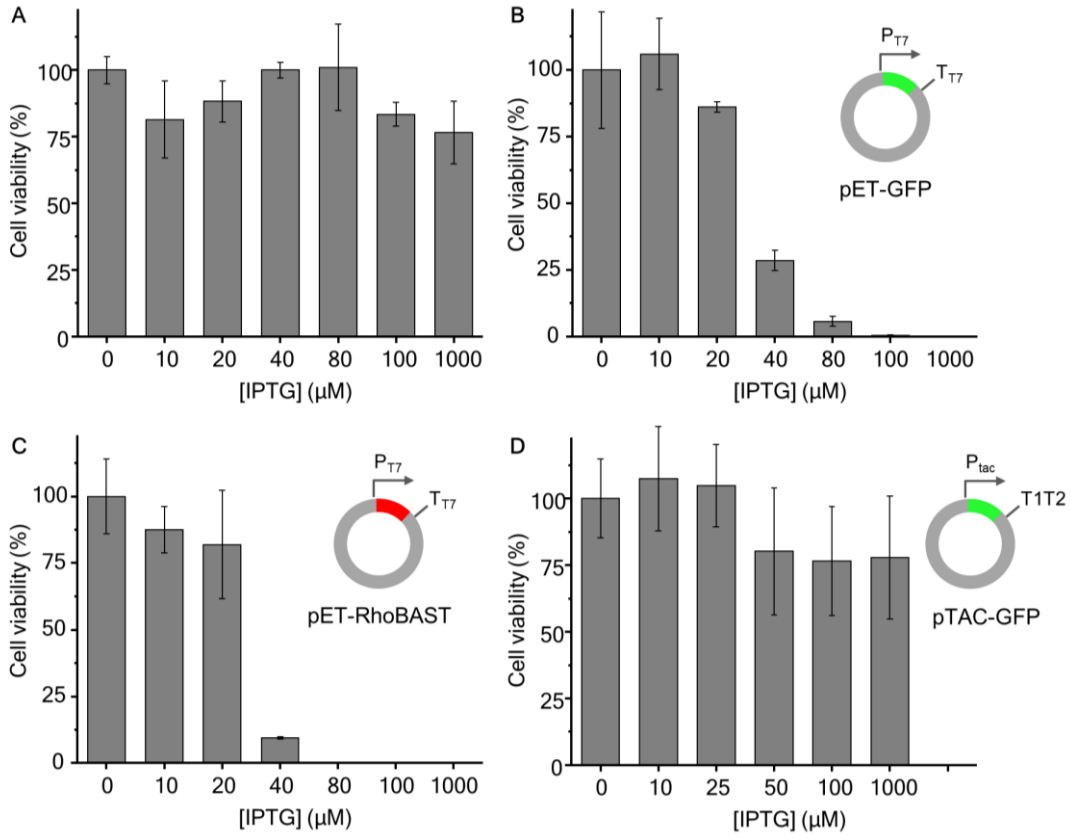


Figure 62. Assessment of bacterial viability with different expression vectors. The BL21 cells transformed with no (A), pET-GFP (B), pET-RhoBAST (C), and DH5α cells transformed with pTAC-GFP (D) were treated with varying concentrations of IPTG (0 - 1 mM). The viabilities determined by colony counting are shown

Changing the promoter from T7 to tac and shifting the host strain from BL21 to DH5α cells significantly rescued the cells (Figure 62D). However, the flow cytometry analysis suggested that DH5α cells transformed with the pTAC expression vector had a lower aptamer expression level than BL21 cells transformed with the pET expression vector (data not shown). Considering the importance of a high expressing level to the FACS-based aptamer selection, we decided to overcome the potential IPTG-induced cytotoxicity of pET-transformed BL21 cells by performing bacteria fixation. The fixation of bacteria has no adverse influence on the aptamer folding or fluorescence activation of the aptamer to its cognate ligand[65]. The fixed aptamer-expressing bacteria could be sorted based on the fluorescence signal, and the enriched aptamer sequences could be fully recovered by PCR amplification of the sorted cells.

3.2.1.4 Mock aptamer-expressing bacteria sorting

Although aptamer-expressing bacteria exhibited good cellular contrast by optical microscopy[56, 65], their fluorescence performances in FACS had not been well studied. A model system, RhoBAST-TMR-DN, was selected to test if the aptamer-expressing bacteria could initiate sufficient fluorescence to differentiate them

from negative control cells. RhoBAST is a reported rhodamine-binding aptamer with a 26-fold fluorescence turn-on and a K_D of 15 nM towards TMR-DN *in vitro* (chapter 1.2.3.5)[59]. RhoBAST aptamer was embedded in a tRNA scaffold within a pET vector and expressed in BL21 *E. coli* cells. A tRNA expression vector was used as a negative control. Transformants treated with varying IPTG concentrations were incubated with different amounts of TMR-DN[f] and then analyzed in the cell sorter. When comparing the RhoBAST+ and RhoBAST- cells, there was, unfortunately, no clear difference in fluorescence distribution of the two populations (Figure 63). This was presumably due to the moderate signal-to-background ratio (several to tens of fold) of the RhoBAST system in cells that made it hard to be distinguished by FACS[59]. For cells stained with the same amount of TMR-DN, the same gate was applied based on the respective negative control where 99.9% of the RhoBAST- cells were kept in Q3. An increase in the percentage of highly bright cells (Q2) from 0.1% to 2.9% (20 nM TMR-DN)/4.2% (100 nM TMR-DN)/1.6% (500 nM TMR-DN)/1.9% (1 µM TMR-DN) was observed when IPTG concentration arose from 0 to 100 µM. The elevated aptamer expression level driven by IPTG promoted the formation of the fluorescent aptamer-ligand complexes. Further increase of the IPTG concentration to 1 mM caused a slight fluorescence decrease, which might be owing to the IPTG induced cytotoxicity. Meanwhile, increasing the dye concentration from 20 to 100 nM resulted a higher mean cellular fluorescence as well as a elevated bright cell fraction for each IPTG-treated group (0.3% to 0.4% for 0 µM IPTG, 1.5% to 1.7% for 20 µM IPTG, 1.6% to 3.5% for 40 µM IPTG, 2.9% to 4.2% for 100 µM IPTG, and 1.5% to 2.6% for 1 mM IPTG) due to the enhanced RNA-fluorophore complex formation. Nevertheless, further increase of the TMR-DN concentration (500 and 1000 nM) led to an escalated background signal from the unbound dye, which offset fluorescence enhancement from the complex formation.

[f] TMR-DN was synthesized and kindly provided by Dr. Mural Sunbu.

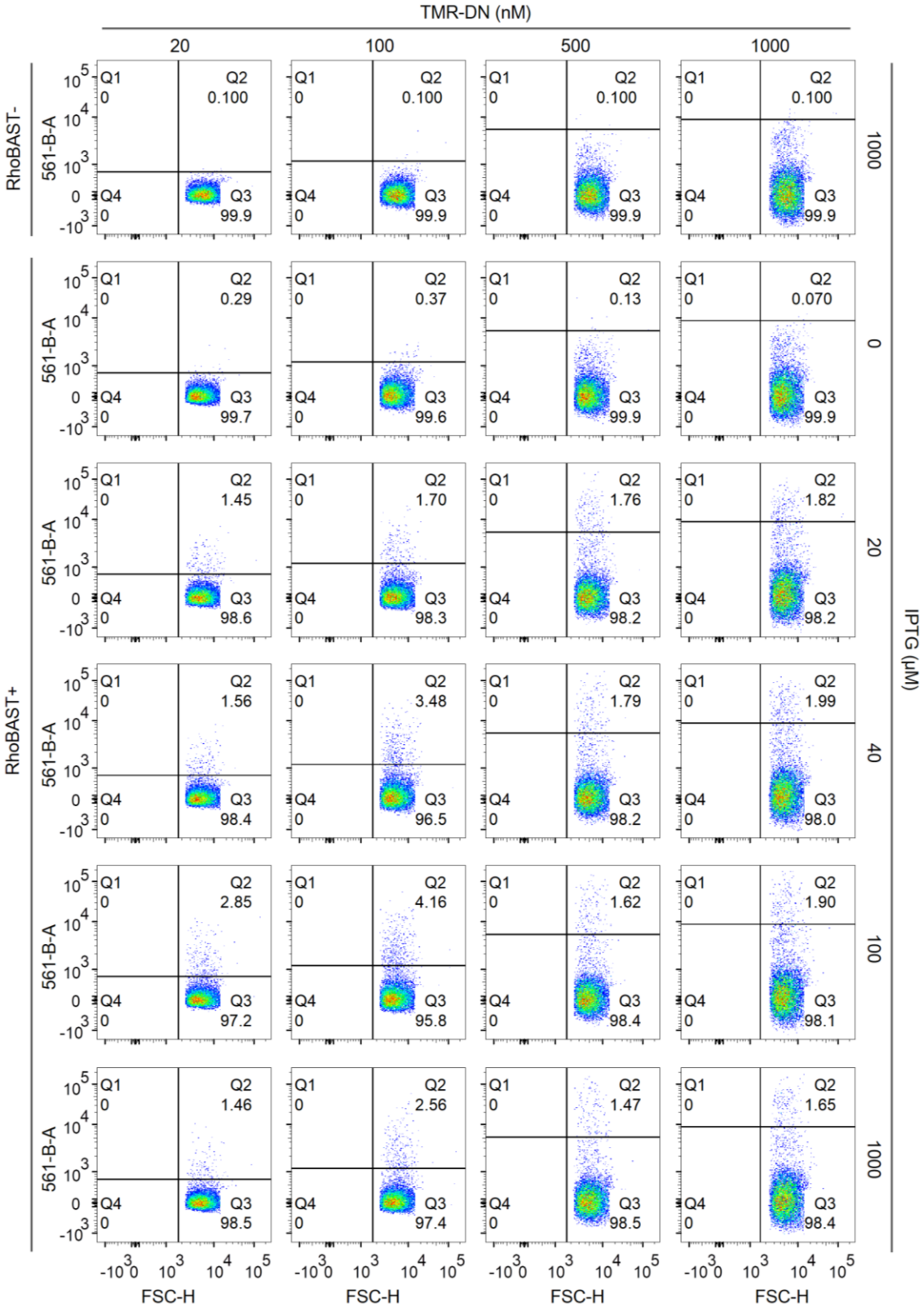


Figure 63. Flow cytometry analysis of aptamer-expressing cells. pET-RhoBAST plasmid and pET-tRNA plasmid as control were transformed into E. coli cells. The cells were treated with different amounts of ITPG (0 - 1 mM), followed by fixation in 1% paraformaldehyde (PFA). The fixed cells were incubated with varying TMR-DN concentrations (20 – 1000 nM) before subjecting to the cell sorter. The cells were excited

with a 561 nm laser, and the scatter plots of FSC vs. fluorescence are shown. Each dot represents one event, and a total of 10^4 events are displayed. The fractions of each quadrant are indicated. The same gate was applied within each TMR-DN group, making 0.1% of the control population display in Q2 while the rest exist in Q3.

To further confirm the capability of FACS for efficient isolation of interesting aptamer-expressing bacteria, mock sorting was conducted. The pretreated RhoBAST+ and RhoBAST- cells were mixed in a 1 : 1, 1 : 10, or 1 : 100 manner, and the highly bright cells were collected (Figure 64A,B). PCR analysis of the output (sorted) cells showed an enhanced band coming from the RhoBAST+ cells and a diminished band coming from the RhoBAST- cells relative to the input (premix) cells, indicating the successful enrichment of RhoBAST+ cells after sorting (Figure 64C). And the enrichment was observed even if the target cells consisted of only 1% of the total population.

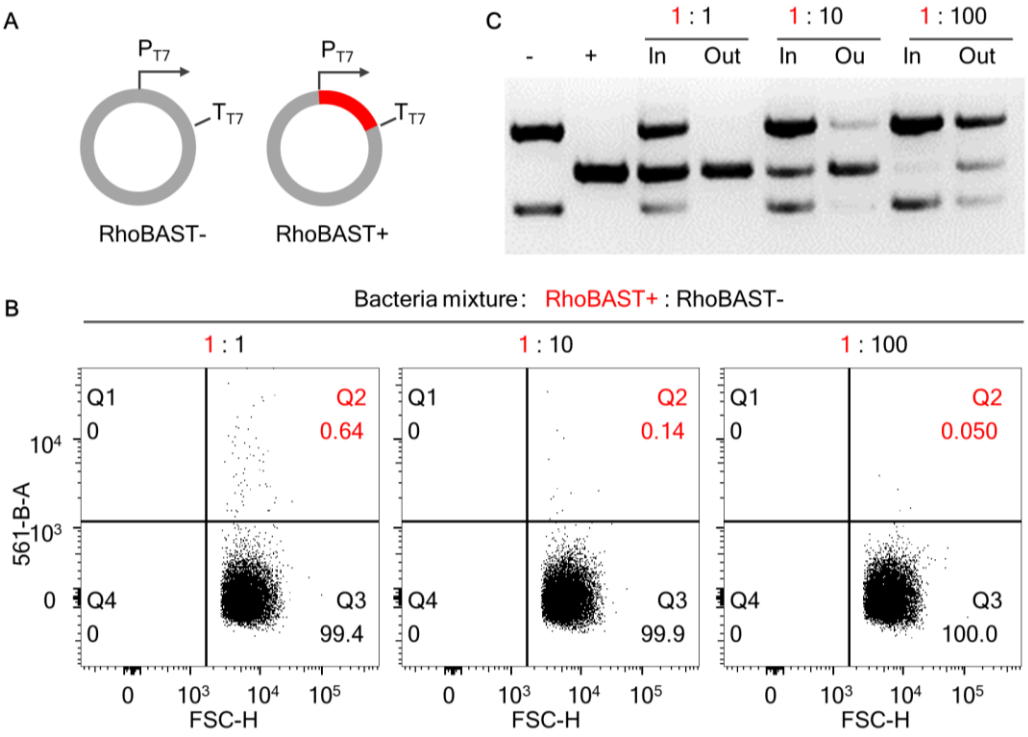


Figure 64. Mock sorting of aptamer-expressing bacteria cells. (A) Illustration of a control (RhoBAST-) vector and a RhoBAST-expressing (RhoBAST+) vector. (B) Flow cytometry analysis of aptamer-expressing bacteria. pET-RhoBAST plasmid and pET-tRNA plasmid as control were transformed into E. coli cells. The cells were treated with 100 μM of ITPG followed by fixation in 1% PFA. The fixed cells were incubated with 100 nM of TMR-DN, mixed at a ratio of 1 : 1, 1 : 10, and 1 : 100 (RhoBAST+ : RhoBAST-), respectively, and subjected to the cell sorter. The cells were excited with a 561 nm laser, and the scatter plots of FSC vs. fluorescence are shown. Each dot represents one event, and a total of 10^4 events are displayed. The same sorting gate was applied to each cell mixture, and the highly fluorescent cells in Q2 were collected. (C) PCR analysis of mock aptamer-expressing bacteria sorting result. Same quantities of RhoBAST- cells, RhoBAST+ cells, cell mixtures (input), and sorted cells (output) were used for PCR amplification. The PCR products were analyzed in an agarose gel.

Fluorescence-activated bacteria sorting for evolving genetically encoded RNA aptamers was extremely difficult due to the small size of bacteria and the intrinsic

relatively low "signal" to "noise" ratio of aptamer-expressing bacteria. By systematical investigation, we found out several parameters that were critical to accomplish this task besides the conventional instrument set up and sample preparation principles for bacterial sorting. (1) A pET vector in BL21 strain to ensure maximal aptamer expression is required. (2) Fixed bacteria are preferred for the selection process, avoiding the loss of diversity due to the different surviving rates. (3) To sufficiently enrich the interesting subpopulation, a proper gating strategy guided by the negative control is requested. In conclusion, the FACS-based platform combined with a bacterial over-expression system showed great potential for the directed evolution of aptamer. Real aptamer selection in bacteria would be performed and discussed in the following texts.

3.2.2 SiR-binding fluorescent light-up aptamer selection via fluorescence-activated bacteria sorting

FACS-based aptamer selection is an ideal technique to enrich aptamer sequences that specifically bind and activate the fluorescence of the spirocyclization-based fluorogenic dyes. Using SiR as a model ligand, we carried out the evolution of SiR-binding aptamer by FACS based on the optimized sorting protocol (chapter 3.2.1).

3.2.2.1 FACS-based aptamer selection scheme

A FACS-based aptamer selection cycle for spirocyclization-based fluorogenic ligands is illustrated (Figure 65). (I) Cloning. The aptamer sequences (i.e., DNA library) are cloned into a pET expression vector. The DNA library design is crucial. It should contain a good quantity of affinity binders. The required sorting time and cycles are substantially diminished, and a better signal separation is achieved for the starting rounds. (II) Transformation. Next, the generated plasmid library is transformed into BL21 *E. coli* cells via electroporation. Good transformation efficiency is required, which ensures adequate coverage of the diversity of the aptamer library. (III) Sample preparation. The rescued bacterial cells after electroporation are cultivated, and the aptamer expression is induced and controlled by the addition of ITPG. Maximal aptamer expression is required for the first selection round to secure a good separation between "positive" and "negative" cells. But the expression level could be reduced as selection pressure for directed aptamer evolution. After induction, the aptamer-expressing cells are fixed to sustain the aptamer concentration in bacteria and avoid any loss of interesting cells, which might undergo cytotoxicity or stress-induced cell death during live-cell sorting. Cells with suitable density (10^6/mL) are incubated with the fluorogenic ligand (the probe quantity could be adjusted each round as directed selection pressure) and subjected to the cell sorter. (IV) Sorting. The bacteria are sorted based on the fluorescence signal of individual cells. An appropriate sorting gate is a prerequisite for successful enrichment. For the aptamer selection, the sorting gate is set using the fluorescence signal of the

negative control as a reference and adjusted accordingly for each round to control the selection pressure. The required number of target cells is collected. (V) PCR amplification. The enriched aptamer sequences from the sorted cells are PCR amplified. The obtained DNA pool is used as input for the next round of selection. A total of 3 to 5 rounds of selection are usually sufficient to evolve a fully enriched subpopulation in favor of binding sequences through FACS[129, 136-137].

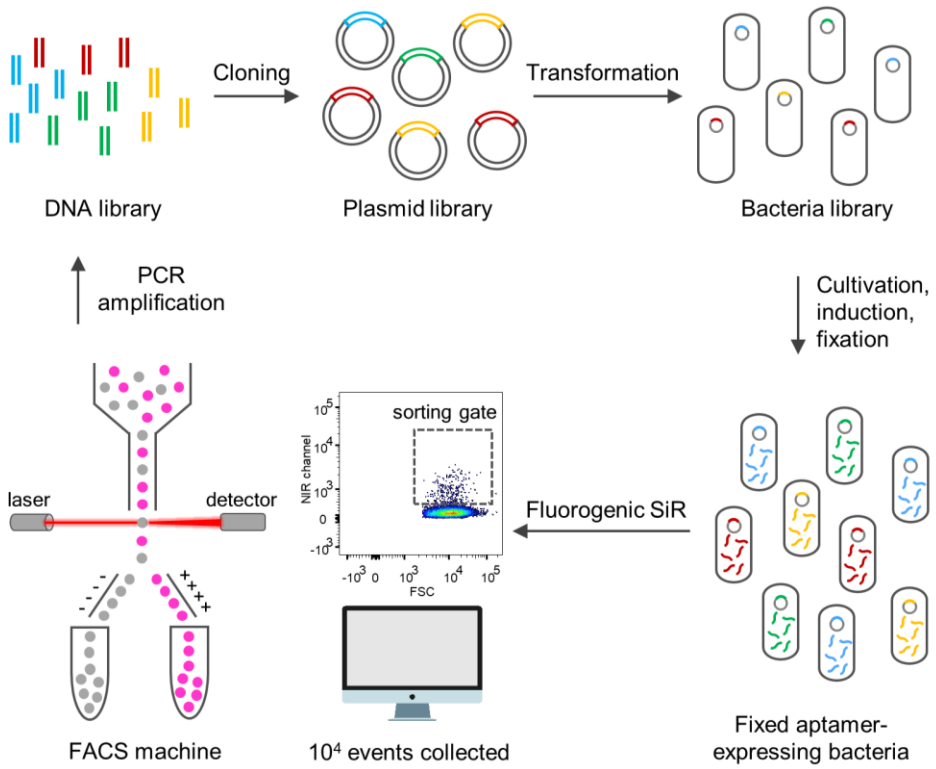


Figure 65. A typical FACS cycle for aptamer selection. The aptamer sequences (DNAs) are cloned into a plasmid vector to generate the recombinant plasmid library. The transformants are cultured, induced, and fixed. The fixed bacteria are incubated with fluorogenic ligand and analyzed on the cell sorter. A proper sorting gate is applied, and highly bright cells are collected. The enriched aptamer sequences are PCR amplified and inserted back to the vector for the next round of selection.

3.2.2.2 Selection of fluorogenic SiR-based ligand

SiR, as a fluorogenic probe, displays an environmentally sensitive NIR fluorescence due to the intramolecular spirocyclization reaction. SiRA binds SiR-PEG_3-NH_2 with a K_D value of 430 nM and a turn-on factor of 7 (Figure 66A)[65]. Additionally, a hydroxymethyl SiR, HMSiR also exhibited a dramatic fluorescence enhancement (> 130-fold) towards SiRA binding but a much worse binding affinity (K_D = 10 µM)[138]. These two SiR derivatives were considered as exciting ligand candidates for the aptamer reselection through FACS. SiRA-expressing cells were incubated with either SiR-PEG_3-NH_2 or HMSiR [g] and analyzed. The cells stained with SiR-PEG_3-NH_2 showed higher fluorescence in

[g] The probes SiR-PEG_3-NH_2 and HMSiR was synthezied and kindly provided by Dr. Regina Wirth.

the NIR channel than that of control cells, while the HMSiR treated cells displayed ultra-low fluorescence signal in both groups (Figure 66B,C). Even if we increased the concentration of HMSiR from 1 μM to 10 μM, no apparent difference between SiRA-expressing and control cells was observed, although HMSiR exhibited much higher fluorogenicity *in vitro* towards SiRA (data not shown). This confirmed our assumption that a good quantity of high-affinity binders is required to initiate a detectable fluorescence difference by the FACS analyzer. Thus, SiR-PEG$_3$-NH$_2$ was chosen as the ligand for the aptamer selection. HMSiR was eliminated due to the unsatisfying fluorescence behavior in FACS.

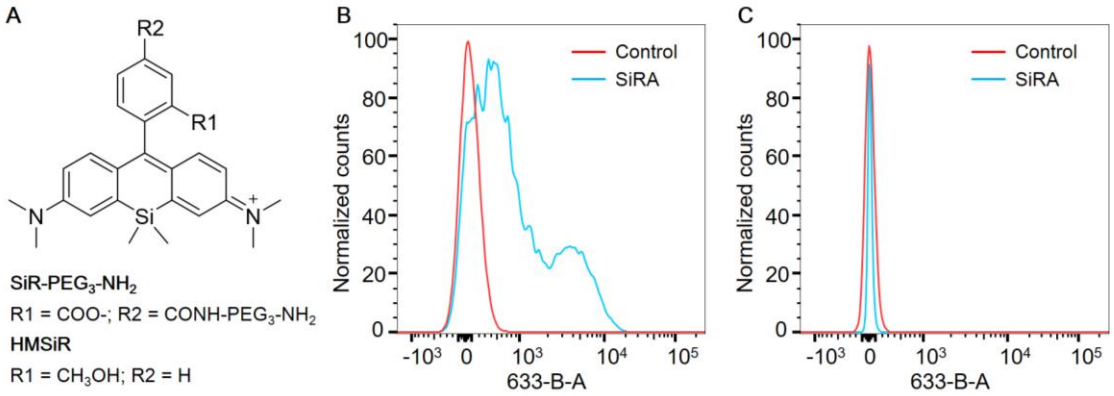


Figure 66. The fluorescence performance of SiR derivatives in SiRA-expressing cells. (A) The chemical structures of SiR-PEG$_3$-NH$_2$ and HMSiR. Fluorescence analysis of E. coli cells stained with SiR-PEG$_3$-NH$_2$ (B) or HMSiR (C). Bacteria were transformed with either pET-tRNA-SiRA[h] or pET-tRNA plasmid. The RNA expression was induced by the addition of IPTG (1 mM). The cells were then fixed in 1% PFA, incubated with 1 μM of SiR-PEG$_3$-NH$_2$ or HMSiR, and analyzed in a cell sorter with a 633 nm excitation laser.

3.2.2.3 Library preparation

The aptamer sequences used as input for aptamer selection are crucial. To evolve new SiRA variants, we used aptamer libraries derived from two origins: the pre-enriched RNA pools against SiR and a doped SiRA library. They allowed a higher abundance of SiR-PEG$_3$-NH$_2$-binding sequences at the beginning stage than a library with completely random sequences.

The enriched pool from SELEX

SiRA was evolved through an *in vitro* aptamer selection against SiR. However, SELEX was affinity oriented. Some good aptamer candidates with high brightness might be lost during the selection due to the less competitive binding affinities[43, 48]. Besides, the Sanger sequencing of 96 colonies from the final pool did not cover the full diversity of the pool. Some good binders might be missed owing to the small sequencing sample size. Therefore, the output DNA pools of

[h] The *pET-tRNA-SiRA* plasmid was cloned and kindly provided by Dr. Regina Wirth.

SELEX rounds 7 and 14, representing the middle and the end of the selection, were picked[i]. The library design for SiR-binding aptamer selection was the same as that for **BeCA** selection (chapter 3.1.2.3). Two restriction sites were added to 5′ and 3′ ends of the aptamer sequences, respectively, to insert the enriched pools to a plasmid backbone (Table 7).

DNA library derived from SELEX pools	5′ - AAA AAA GTC GAC GGA GCT CAG CCT TCA CTG C N_{26} CTG CTT CGG CAG N_{26} GGC ACC ACG GTC GGA TCC AC GAC GTC TTT TTT– 3′

Table 7. The sequence of the pre-enriched DNA pools from SELEX against SiR. SalI and AatII sequences are indicated in blue and green, respectively.

Doped SiRA library

Another partially randomized (doped) library based on the SiRA template sequence was designed and chemically synthesized to improve the reselection outcome. The biased pool could optimize the previously isolated SiRA by retaining the critical binding motifs while exploring the surrounding "sequence space"[139]. A doping rate of 9%, which means an average of 4.5 mutations/SiRA template (0.9% probability mutation/position × 50 positions = 4.5 mutations), was used in the library design. The doped library was pre-installed with two restriction sites flanking at the doped sequence.

Doped SiRA	5′ – G CTC AGC CTT CAC TGC GTC GAC GGT (N3)(N3)(N2) (N2)(N1)(N2) (N2)(N3)(N3) (N3)(N4)(N4) (N4)(N3)(N1) (N1)(N1)(N1) (N2)(N2)(N4) (N3)(N3)(N2) (N4)(N3)(N2) (N4)(N4)(N2) (N3)(N3)(N2) (N1)(N3)(N4) (N4)(N3)(N4) (N1)(N4)(N2) (N2)(N4)(N4) (N4)(N3)(N3) (N2)(N2) GAC GTC GAT GGC ACC ACG GTC GGA T – 3′

Table 8. The sequence of the 9% doped SiRA library. N represents a nucleotide with the probabilities in the order of ACGT (N1: 91030303; N2: 03910303; N3: 03039103; N4: 03030391). SalI and AatII sequences are indicated in blue and green, respectively.

Expression vector preparation

A proper expression vector was needed to generate a recombinant aptamer-expressing plasmid library. The use of a scaffold RNA such as tRNA or F30 to accommodate recombinant RNA is proven to enhance the folding of RNA insert due to the high structural stability[123, 140]. Therefore, we cloned a scaffold tRNA containing two restriction sites at the anticodon region into a pET backbone. The generated *pET-tRNA-2* was used as the vector to introduce the aptamer sequences into the bacterial cells (Table 9).

[i] The enriched DNA pools from SELEX against SiR were generated and kindly provided by Dr. Regina Wirth

pET-tRNA-2	5′ – …GCC CGG ATA GCT CAG TCG GTA GAG CAG C GTC GAC GGT TTT GAC GTC GAT GCG GGT CCA GGG TTC AAG TCC CTG TTC GGG CGC CA… – 3′

Table 9. The sequence of tRNA-2 in the pET-tRNA-2 plasmid. The tRNA scaffold, Sal I, and AatII sequences are indicated in orange, blue, and green, respectively

Recombinant plasmid library preparation

The aptamer libraries were then inserted to the center of the tRNA sequence within *pET-tRNA-2* after double-digestion and sticky-end ligation. (Figure 67) We named the three plasmid libraries L1 (derived from SELEX round 7), L2 (derived from SELEX round 14), and L3 (derived from doped SiRA), respectively.

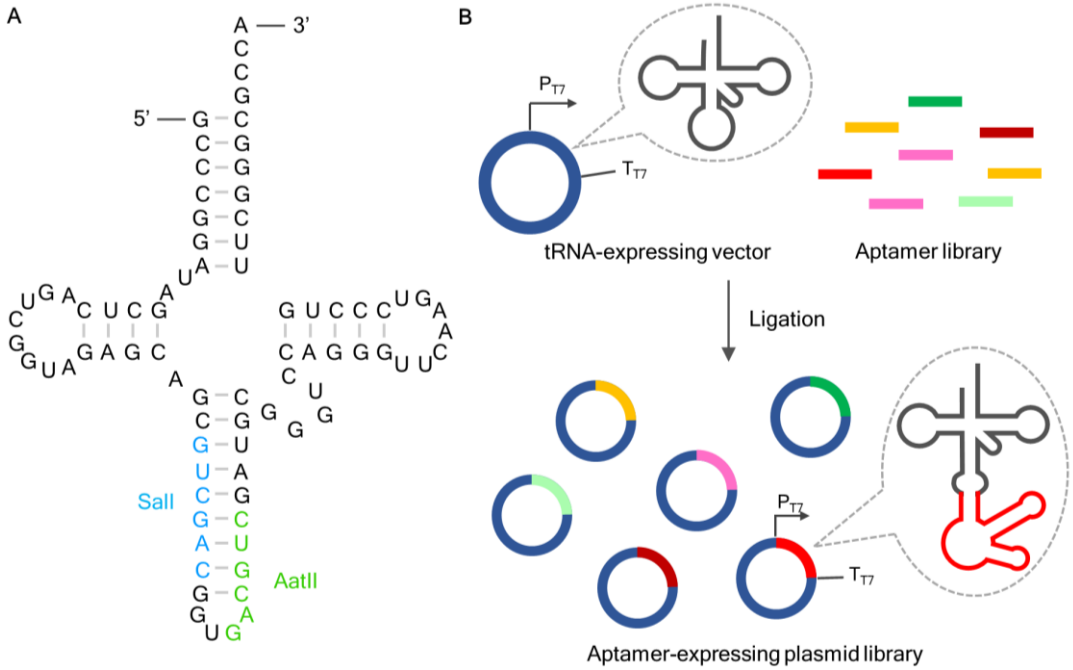


Figure 67. Recombinant plasmid library preparation. (A) The sequence and secondary structure of a recombinant tRNA. The corresponding RNA sequences of SalI and AatII restriction sites were labeled in blue and green, respectively. (B) Preparation of an aptamer-expressing plasmid library. The aptamer library was cloned into a tRNA-expressing vector to generate an aptamer-expressing plasmid library. The expression products, tRNA and tRNA-aptamer hybrid (circled in gray), are shown. P represents promoter while T stands for terminator.

3.2.2.4 Fluorescence-activated bacteria sorting

FACS-based aptamer screening was conducted following the above-described scheme to isolate the highly fluorescent aptamer-expressing bacteria from the bacterial library in the presence of SiR-PEG_3-NH_2 (Figure 65). The selection started with transforming the plasmid libraries (L1, L2, and L3) into *E. coli* cells. Transformation efficiency of ~ 10^5 colony-forming unit (CFU)/μg for L1 and L2 and ~ 10^6 CFU/μg for L3 was achieved. After cultivation, induction, and fixation, the aptamer-expressing cells were fixed and incubated with SiR-PEG_3-NH_2.

Appropriate sorting gates, which were set above the respective negative controls (i.e., tRNA-expressing cells), were applied, and a total of 10^4 highly bright cells were collected for each library group. Three iterative selection rounds were performed.

The fluorescence behavior of the aptamer-expressing cells was recorded for each selection round. A separation between "bright" and "silent" cells was achieved by applying a gate using the control cells as a reference. Analysis of the scatter plots showed the percentage of bright cells in Q2 grew from 1.2 to 3.6 for L1, 0.3 to 4.1 for L2, and 0.3 to 1.5 for L3 after three selection rounds when incubated with 1 µM of SiR-PEG$_3$-NH$_2$, indicating a successful screening and enrichment of FLAPs using FACS (Figure 68). In parallel with a "normal" (1 µM) probe concentration used in round 3, a low SiR-PEG$_3$-NH$_2$ concentration (200 nM) was applied in round 3-L as exogenously applied selection pressure. As expected, the percentage of bright cells went down to 1.3, 1.5, and 0.4 for L1, L2, and L3, respectively, due to the loss of low-affinity binders, but still significantly higher than that of the control cells where 0.05 was set. The aptamer sequences amplified from the sorted cells incubated with either a "normal" (round 3) or a low (round 3-L) probe concentration were sent for sequencing.

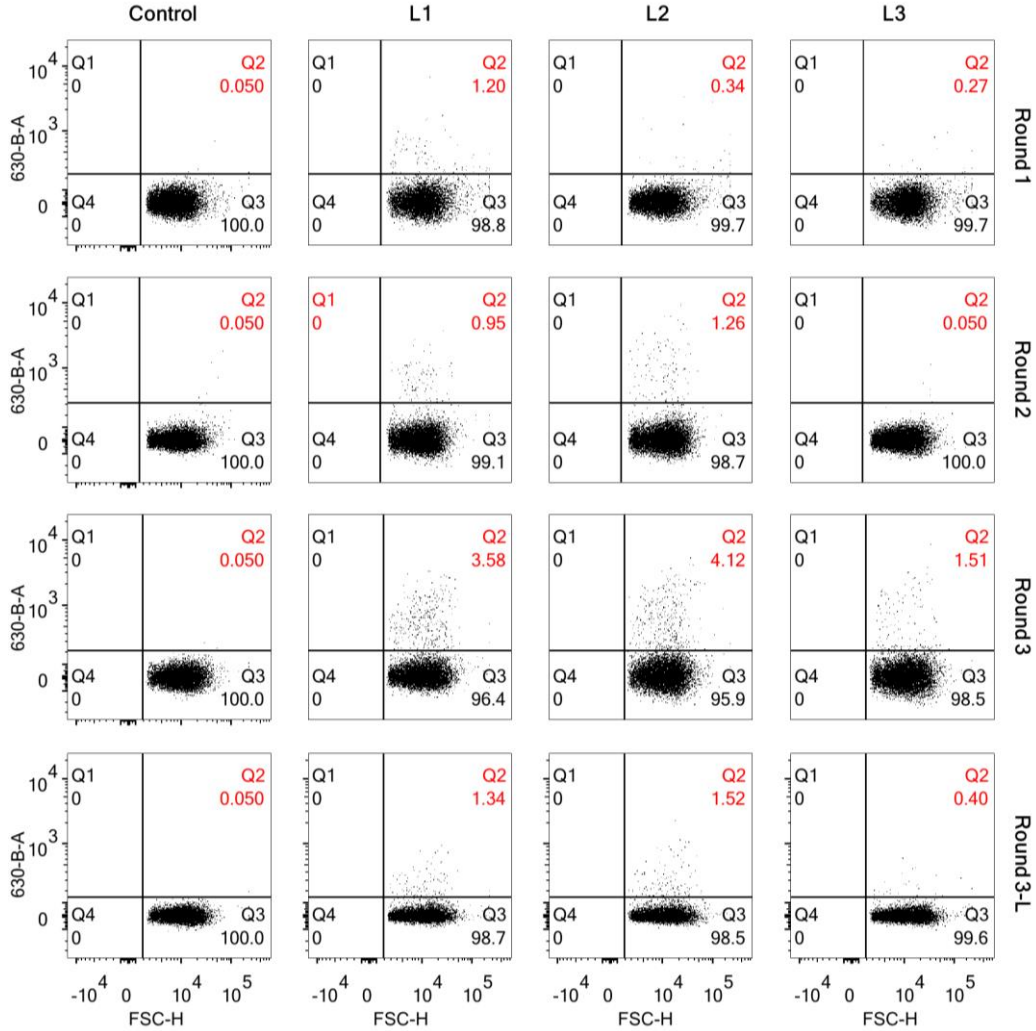


Figure 68. Flow cytometry analysis of E. Coli cells from each sorting round. The cells were excited with a 633 nm laser, and the scatter plots of FSC vs. fluorescence are shown. Each dot represents one event, and a total of 10^4 events are displayed. The fractions of each quadrant are indicated. The soring gate for each round was applied, where 0.05% of the control population was in Q2 while the rest was in Q3. SiR-PEG_3-NH_2, 1 μM and 200 nM, respectively, were used for rounds 1-3 and round 3-L.

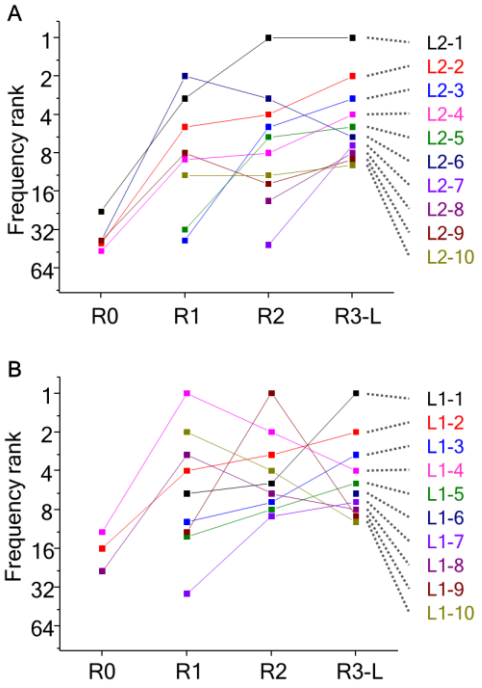
3.2.3 Identification of SiR-binding motif

The isolated aptamer sequences by FACS were disclosed by sequencing. To identify the SiR-binding motifs, *in vitro* screening, characterization, and truncation of the highly enriched sequences were carried out.

3.2.3.1 Sequence analysis

The enriched DNA pools from each round (R0 to R3) for L1 and L2 were sent for next-generation sequencing (NGS)[j]. The top 10 sequences (by frequency, see appendix 1) from the last selection round, R3-L, for both libraries were analyzed. Enrichment was observed for the majority of these top 10 sequences during the course of selection, indicating that these variants outcompeted others in the fluorescence-activated sorting and were promising candidates as SiR-binding FLAPs in a cellular milieu (Figure 69). Interestingly, a few sequences, including L2-6, L1-4, L1-8, and L1-10, showed a stepwise ranking decline after R1, which could be reasoned to the increased selection pressure. It is worth noting that there are apparent ranking differences between R3 and R3-L for both L1 and L2, even though the selection of R3 and R3-L started with the same aptamer library. We assumed that reducing ligand concentration (1 μM to 200 nM) changed the direction of aptamer enrichment significantly. Only the sequences with higher binding affinity could form stable fluorescent complexes at the lower ligand concentration and were therefore enriched. This finding also confirmed the capability of FACS for directed aptamer evolution in favor of both brightness and binding affinity.

[j] The NGS were performed and analyzed by Dr. Yaqing Zhang.



L2	R0	R1	R2	R3-L	R3
Frequency rank	23	3	1	1	1
	41	5	4	2	2
	-	39	5	3	6
	47	9	8	4	5
	-	32	6	5	4
	39	2	3	6	3
	-	-	42	7	10
	-	-	19	8	11
	39	8	14	9	7
	-	12	12	10	12

L1	R0	R1	R2	R3-L	R3
Frequency rank	-	6	5	1	11
	16	4	3	2	4
	-	10	7	3	15
	12	1	2	4	6
	-	13	8	5	19
	-	-	-	6	1
	-	36	9	7	28
	24	3	6	8	16
	-	12	1	9	2
	-	2	4	10	3

Figure 69. The evolution paths of the highly enriched sequences through the selection cycles. The top 10 sequences ranked by appearing frequency in R3-L for L2 (A) and L1 (B) and their ranking positions in previous rounds are shown.

Moreover, the analysis of the top 10 sequences for both L1 and L2 from R3-L showed that six sequences were appearing in both libraries. Thus, a total of 14 unique sequences were subjected to alignment to find common motifs. Five families were identified based on the conserved sequences and clustered with different colors (Figure 70).

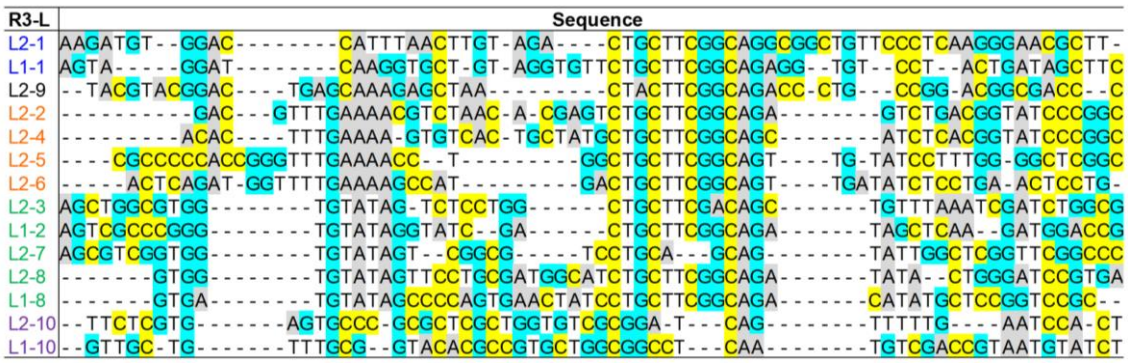


Figure 70. Alignment of the highly enriched sequencers from R3-L using MultAlin[111]. The primer-binding sequences at the 5′ and 3′ ends of each sequence are not shown. The sequences were clustered into five families indicated with the different colors on their names (blue, black, orange, green, and purple, respectively). The alignment was performed using a symbol comparison table of altDNA, a gap weight of 8, and a gap length weight of 0.

Besides NGS, we also performed Sanger sequencing for the isolated aptamer sequences. The output DNA pools of L1, L2, and L3 from R3 and R3-L were cloned into a pDisplay vector and transformed into *E. coli* cells, respectively.

Thirty-two colonies were picked for each pool and sent for sequencing (result see appendix 2). As expected, the revealed sequences for L1 and L2 were already covered by NGS data. For L3, 18 and 11 unique sequences for R3 and R3-L were obtained, respectively. However, larger sample size and further analysis are needed to determine which mutations are essential for the optimization of SiRA.

3.2.3.2 Sequence characterization

We next examined the fluorescent response of SiR-PEG_3-NH_2 towards the 14 highly enriched sequences from L1 and L2. The fluorescence of the RNA-SiR-PEG_3-NH_2 complexes was recorded at their emission maxima. The critical parameters for assessing a FLAP, dissociate constant and fluorescence turn-on, were obtained from the fitted fluorescence curve and utilized (Figure 71 and Table 10).

L2-1 in class I showed a 4.4-fold fluorescence turn-on and the best binding affinity (K_D = 25 nM) among all the sequences, which is 17-fold better than that of SiRA. A ~ 10 nm red-shift of both the excitation and emission maxima relative to the unbound fluorophore was observed, indicating the interaction between the fluorophore and RNA. Interestingly, L1-1 from the same class also exhibited a good binding affinity (K_D = 141 nM) but displayed slow binding kinetics. It took more than 30 min for the RNA-fluorophore complex to reach the fluorescence equilibrium. This was done within sec for the other sequences (Figure 72). A reshaping of the RNA to accommodate the fluorophore might be involved in the "induced-fit" process. L2-2 is the best candidate in class II, showing a K_D value of 208 nM and a 5.3-fold fluorescence enhancement. Notably, the parent sequence of SiRA, L2-5, was also present in this class. One unique feature of this class was that they all inherited the high brightness of SiRA. On average, they were twice brighter than the rest aptamers from other classes when the aptamer-fluorophore complexes were excited at 649 nm (Table 10). As the most abundant cluster, class III consists of five sequences with bathochromic shifted (~ 15 nm) excitation and emission. Among them, L2-8, as the best variant, displayed a nanomolar K_D (53 nM) and a 5.6-fold fluorescence turn-on. Unfortunately, the three variants from class IV and V did not show promising fluorescence responses *in vitro*. They might fold differently in the presence of a tRNA scaffold in a cellular environment. In conclusion, *in vitro* screening and characterization of the highly enriched sequences disclosed three SiR-binding aptamer families with improved affinity.

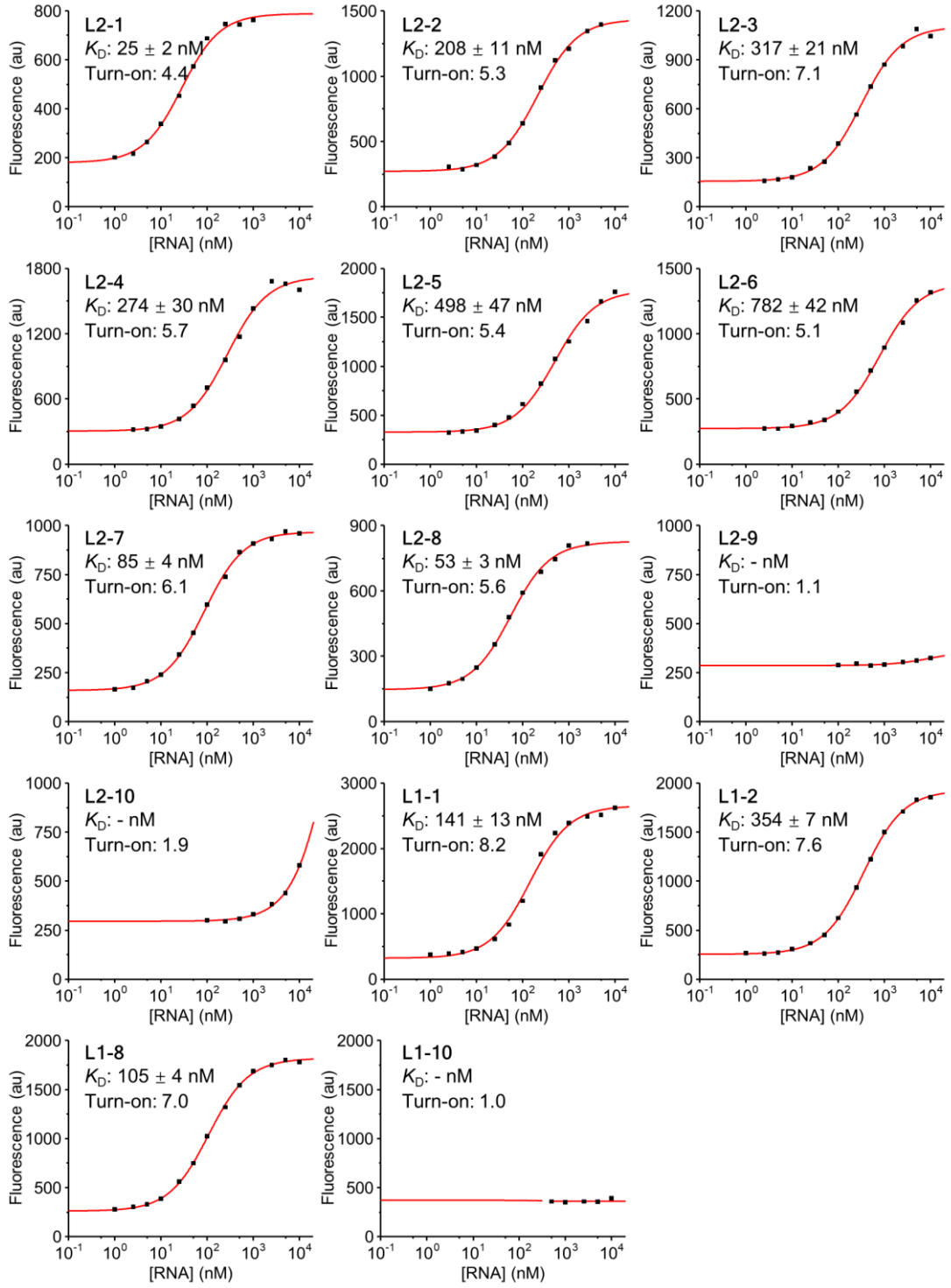


Figure 71. Characterization of the fourteen highly enriched RNA sequences. SiR-PEG$_3$-NH$_2$ (5 nM) was titrated with different amounts of RNA (0 - 10 µM). The K_D values and fluorescence turn-on factors (F_{max}/F_{min}) were determined from the fitted curve. The RNA-fluorophore complexes were excited at their excitation maxima (± 5 nm) in 1×ASB buffer containing 20 mM HEPES (pH 7.4), 5 mM MgCl$_2$, 125 mM KCl, and 0.05% tween 20.

Name	K_D (nM)	Turn-on[a]	Turn-on[b]	λ_{ex} (nm)	λ_{em} (nm)	Class
L2-1	25	4.4	2.5	659	674	I
L2-2	208	5.3	5.3	649	661	II
L2-3	317	7.1	2.0	664	679	III
L2-4	274	5.7	5.7	649	662	II
L2-5	498	5.4	5.4	649	660	II
L2-6	782	5.1	5.1	649	661	II
L2-7	85	6.1	2.4	662	676	III
L2-8	53	5.6	2.3	664	676	III
L2-9	-	1.1	1.0	653	665	V
L2-10	-	1.9	1.9	649	660	IV
L1-1	141	8.2	2.9	661	674	I
L1-2	354	7.6	1.9	664	679	III
L1-8	105	7.0	1.8	664	679	III
L1-10	-	1.0	1.0	649	662	IV

Table 10. The spectroscopic characteristics of the fourteen highly enriched RNA sequences. Turn-on[a]: obtained from the K_D curve, where the RNA-fluorophore complexes were excited at their excitation maxima (± 5 nm). Turn-on[b]: division of the fluorophore's maximal fluorescence intensities in the presence or absence of RNA (10 μM) when excited at 649 nm ± 5 nm.

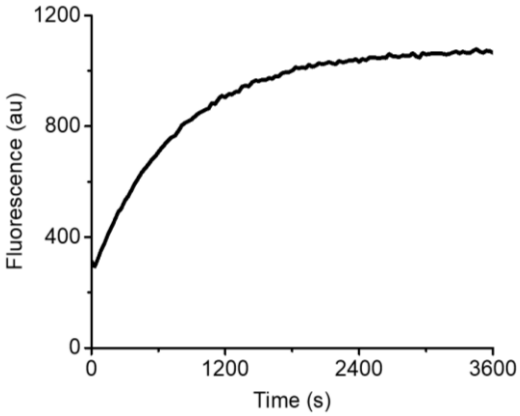


Figure 72. The binding kinetics of L1-1. L1-1 (1 μM) was incubated with SiR-PEG$_3$-NH$_2$ (5 nM) in 1×ASB buffer containing 20 mM HEPES (pH 7.4), 5 mM MgCl$_2$, 125 mM KCl, and 0.05% tween 20. The time-dependent fluorescence was recorded every 30 s for 1 h with an excitation of 661 ± 5 nm immediately after mixing.

3.2.3.3 Truncation

To generate a minimal aptamer motif for intracellular RNA imaging, the best performers from each class (I, II, and III), i.e., L2-1, L2-2, and L2-8, were selected for truncation study.

Based on the predicted secondary structure of L2-2, truncation variants were designed and assayed. Sequential removal of the terminal stem-loops (L2-2-1), truncation of the centrally located stem (L2-2-2), and fixation of the U_4-G_{51} pair

by C-G (L2-2-3) did not affect the binding affinity while slightly increased the fluorescence turn-on (Figure 73). The 53-nt L2-2-3 (K_D = 192 nM, turn-on = 5.5-fold) was determined as the minimal binding pocket for the ligand SiR-PEG$_3$-NH$_2$. Although displaying a similar fluorescence turn-on and brightness to SiRA upon fluorophore binding, this new aptamer showed a 2.2-fold better binding affinity, which indicates a better cellular fluorescence performance of the new aptamer in living cells.

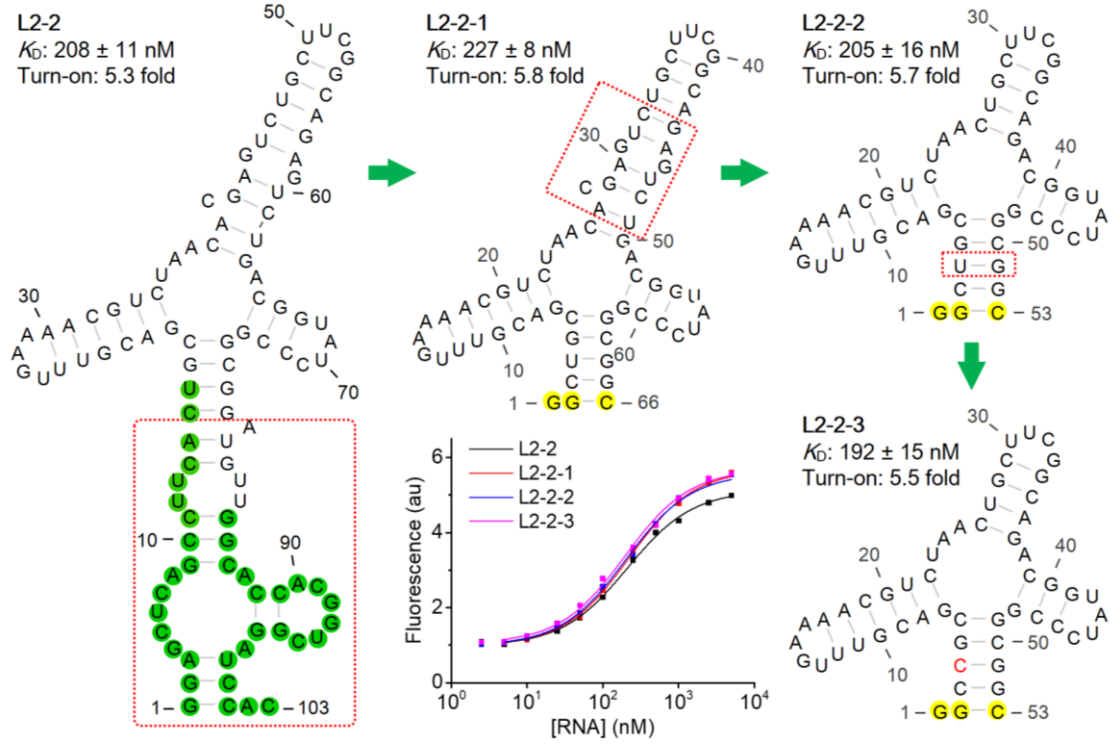


Figure 73. Truncation results of L2-2. The predicted secondary structures of L2-2 and its truncated versions L2-2-1, L2-2-2, and L2-2-3 are shown. The primer-binding regions are indicated in green, while the introduced nucleotides for transcription efficiency and stem stabilization are shown in yellow. The modification sites are boxed in red. The fluorescence characterization of each variant is shown. SiR-PEG$_3$-NH$_2$ (5 nM) was titrated with different amounts of RNA (0 - 5 μM). The K_D values and fluorescence turn-on factors (F_{max}/F_{min}) were determined from the fitted curve. The RNA-fluorophore complexes were excited at 649 ± 5 nm in 1×ASB buffer containing 20 mM HEPES (pH 7.4), 5 mM $MgCl_2$, 125 mM KCl, and 0.05% tween 20.

Although the great binding affinity of L2-1 and L2-8 already enabled them to be great candidates for live-cell RNA imaging, the truncation was conducted to reduce their sequence length further. Three truncation variants were designed by deleting the primer-binding region from 5′ end (-1), 3′ end (-2), or both ends (-3) for L2-1 and L2-8, respectively. Unfortunately, deleting the primer-binding region from either end reduced the affinity remarkably, as indicated by the low fluorescence turn-on (Figure 74). This implied the vital role of the primer-binding sequences in the fluorophore complexation. Further truncation attempts based on the predicted alternative foldings of L2-1 and L2-8 should be performed.

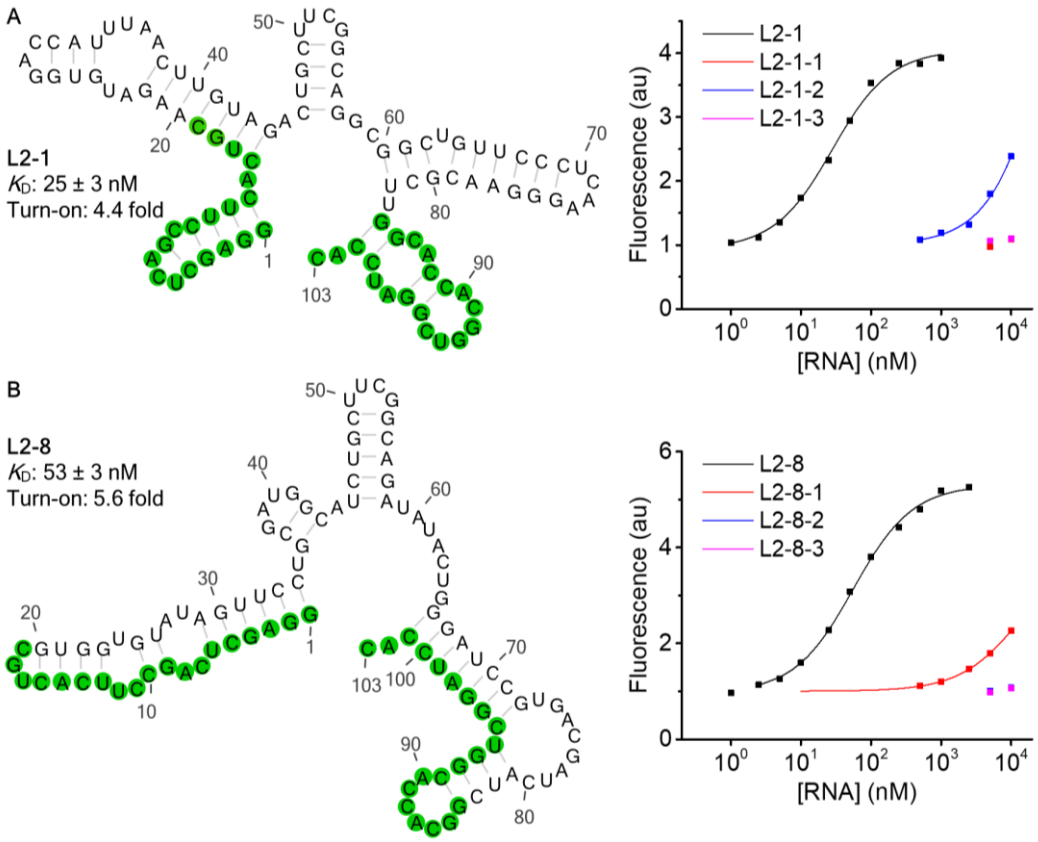


Figure 74. Truncation results of L2-1 (A) and L2-8 (B). The predicted secondary structures of L2-1 and L2-8 are shown. The primer-binding regions are indicated in green. Variants generated by removing primer-binding sites from 5′ end, 3′ end, or both ends were designated as -1, -2, and -3, respectively. The fluorescence characterization of each variant is shown. SiR-PEG$_3$-NH$_2$ (5 nM) was titrated with different amounts of RNA (0 - 5 μM). The RNA-fluorophore complexes were excited at their excitation maximum (± 5 nm) in 1×ASB buffer containing 20 mM HEPES (pH 7.4), 5 mM MgCl$_2$, 125 mM KCl, and 0.05% tween 20.

In conclusion, in this chapter (chapter 3.2), we successfully established a FACS-based aptamer screening platform. We demonstrated its capability in evolving new aptamers with high binding affinity and brightness using a fluorogenic SiR. This new platform solved the bottleneck problems in aptamer evolution for the spirocyclization-based fluorogenic dyes, thus opening up new possibilities for developing numerous new FLAP-based RNA imaging tools.

4 Summary and outlook

The majority of the content described in chapter 4.1.1 was covered in an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.

4.1 Summary

In conclusion, we presented the development of two NIR FLAPs for RNA imaging in this thesis. A summary for each project is provided.

4.1.1 Benzopyrylium-coumarin binding color-shifting near-infrared fluorescent light-up aptamer

In the first project, we presented the development of a color-shifting NIR-fluorescent aptamer-fluorophore module **BeCA-BC6** for live-cell RNA imaging.

4.1.1.1 *In vitro* selection

We first exploited an environmentally sensitive hybrid benzopyrylium-coumarin (BC) fluorophore. It exists in a dynamic equilibrium between a cyan-fluorescent spirocyclic form and a NIR-fluorescent zwitterionic form. A reported (classic) BC fluorophore that predominantly existed in the NIR-fluorescent zwitterionic form in aqueous buffers was synthesized and used as a bait for the *in vitro* aptamer selection (chapter 3.1.1). SELEX started with a combinatorial RNA library containing ~ 3×10^{15} molecules, and 11 iterative selection rounds were successfully proceeded (chapter 3.1.2). Activity screening of the selected RNA sequences revealed eight highly active aptamer candidates. Truncation and mutation studies of the best candidate RNA8 rendered a 38-nt minimal motif **BeCA**, which selectively binds to the NIR-fluorescent BC zwitterion with a nanomolar affinity (chapter 3.1.3). In parallel, we synthesized a series of BC analogs by introducing electron-withdrawing fluorine atoms and electron-deficient amines to BC (chapter 3.1.4). The chemical modifications on BC scaffold fine-tuned the intramolecular zwitterion/spiro-lactone(lactam) ratio. Subsequent fluorescence screening of the promising color-shifting ligands, which display D_{50} values around 50, revealed the best aptamer-fluorophore pair **BeCA-BC6** (chapter 3.1.5). **BC6** exists primarily in the spirocyclic state with an emission maximum of 478 nm, while binding to **BeCA** shifted its emission to peak at 684 nm, representing the most NIR-shifted FLAP in the literature. The emission ratio change (cyan/NIR) of **BC6** upon **BeCA** binding is as high as 15-fold (chapter 3.1.6). Further characterization of **BeCA-BC6** revealed a nanomolar affinity (K_D = 220 nM), good thermal stability, and mono/bi-valent cation independence, making it suitable for RNA imaging at physiological conditions.

4.1.1.2 Live-cell RNA imaging

Next, live-cell imaging was performed using the new color-shifting NIR FLAP **BeCA-BC6**. The spirocyclic **BC6** showed good membrane permeability and low unspecific NIR staining inside the bacterial cells, which are prominent characteristics of the spirocyclization-based probe[73, 82]. By genetically fusing the multimerized **BeCA** tag (eight tandem repeats) to mRNA, a good signal-to-background ratio (~ 6-fold) was achieved (chapter 3.1.7). Thus, **BeCA-BC6** is a valuable addition to the RNA imaging toolbox due to the lack of FLAPs functioning in the NIR region, where cells have much lower absorption (less phototoxicity) and autofluorescence. Moreover, **BeCA-BC6** is the only aptamer-fluorophore pair that allows simultaneous imaging of both the free fluorophore and the complex. The cyan fluorescence from the unbound probe reveals the intracellular probe delivery efficiency and distribution, while the NIR fluorescence indicates the RNA location. We used this feature to obtain ratiometric images of bacteria with different expression levels of **BeCA**-labeled tRNA. In contrast to single-color fluorescence readouts, ratiometric images avoided the problems associated with varying dye uptake, heterogeneous probe distribution, probe instability, cell morphology, and fluctuations in the focal plane. Therefore, the dual-color feature of **BeCA-BC6** allowed us to analyze the expression levels of RNA transcripts more accurately and revealed the transcriptional heterogeneity at the single-cell level. By utilizing aptamer tandem repeats combined with the ratiometric advantages, we anticipate that **BeCA-BC6** could be used to image and analyze the abundance of target RNAs, which could provide new insights into gene expression, regulation, developmental plasticity, and disease diagnostics[141].

4.1.1.3 Sensor development

In the last chapter 3.1.9, a ratiometric sensor platform based on the color-shifting **BeCA-BC6** module was designed. Insertion of a tetracycline-binding aptamer to **BeCA** via a transducer stem generated a tetracycline sensor, showing an EC_{50} of 470 nM, a dynamic range of 2.3-fold, and a detection range of 0.2 - 1 µM *in vitro*. This RNA-based sensor platform offered a new strategy for developing ratiometric sensors for accurate detection of biomolecules.

4.1.2 Silicon rhodamine-binding near-infrared fluorescent light-up aptamer

In the second project, we presented the directed evolution of aptamers for spirocyclization-based fluorogenic dyes using FACS. We first established a standard sorting method for fluorescent aptamer-expressing bacteria, which possess relatively low (several to tens of fold) cellular contrast in flow cytometry (chapter 3.2.1). In addition to proper instrument setup and sample preparation, an aptamer over-expression system (pET expression vector in BL21 Star (DE3)) and appropriate gating strategy are required to isolate and enrich the

populations of interest. Aptamer-expressing bacterial cells were subjected and sorted in a cell sorter starting with the pre-enriched SiR-binding libraries (embedded in a tRNA scaffold; chapter 3.2.2). The selection was directed by NIR fluorescence using a culture medium without additional supplements. Next-generation sequencing of the output DNA pools denoted the evolution paths of the highly enriched sequences through selection (chapter 3.2.3). The promising aptamer candidates were screened in terms of both affinity and fluorogenicity in the presence of SiR, which revealed three SiR-binding sequence families. Among them, a 103-nt variant L2-1 showing the highest binding affinity (K_D = 25 nM) and a moderate turn-on (4.4-fold), and a 53-nt variant L2-2-3 possessing the highest brightness (turn-on: 5.5-fold excited at 649 nm) and a good dissociate constant (K_D = 192 nM). These new aptamers are improved versions of SiRA and are interesting candidates for further intracellular evaluations.

4.2 Outlook

4.2.1 Evolution of the next-generation of color-shifting aptamer-fluorophore module

BeCA-BC6 provided the first example of a color-shifting aptamer-fluorophore module for live-cell RNA labeling and sensing, yet some limitations of the system need to be addressed. **BeCA-BC6** only demonstrated a moderate K_D of 230 nM and the relatively low brightness of both unbound cyan-fluorescent **BC6** ($\varepsilon^* \Phi$ = 1088 $mol^{-1}cm^{-1}$) and NIR-fluorescent complex ($\varepsilon^* \Phi$ = 3689 $mol^{-1}cm^{-1}$), making it difficult for imaging less abundant RNAs in mammalian cells. Optimizations in both aspects of the fluorophore and the aptamer need to be performed to develop the next generation of color-shifting aptamer-fluorophore modules with better affinity, higher brightness, larger dynamic range, and improved thermal stability (Figure 75A).

As described in chapter 3.1.4, **BC7** possessed the highest spirocyclization degree among all BCs in aqueous buffers. Moreover, according to the reported literature, the utility of a stronger intramolecular nucleophile such as acyl sulfamide (X = $NSO_2N(CH_3)_2$) can further regulate BC into the spirocyclic configuration[82, 115]. These highly spirocyclic BCs possess higher brightness in the cyan channel and lower background signal in the NIR channel and yet hold the potential to be efficiently switched to the corresponding NIR-fluorescent zwitterionic form by interacting with the cognate aptamers. They are interesting ligand candidates as they possess a higher cyan brightness, a lower background signal, and potentially a larger dynamic range.

SELEX is an affinity-oriented *in vitro* aptamer selection method. Enrichment of RNA sequences binding to the spirocyclic form instead of zwitterion will be obtained if the spirocyclic BCs are used as selection baits. In this work, we developed a strategy in which an aptamer (i.e., **BeCA**) was first evolved against a zwitterionic BC prototype through SELEX. The chemically modified spirocyclic BCs were then screened for their NIR/cyan emission ratio change towards **BeCA**.

Although **BeCA-BC6** was successfully developed and applied in live-cell imaging, the binding affinity and dynamic range of this color-shifting aptamer-fluorophore module limited by this aptamer evolution method are not satisfying yet. Some highly spirocyclic BCs (e.g., **BC7**) were eliminated after substrate screening due to the affinity loss between the aptamer and a modified ligand.
To evolve color-shifting aptamers for these highly spirocyclic BCs, FACS is an excellent alternative technique since the selection can be set in favor of both affinity and brightness, as demonstrated in chapter 3.2 (Figure 75B,C). To ensure efficient isolation and less required selection cycles, the output DNA pool from SELEX (i.e., SELEX round 11 for **BeCA**) or a doped **BeCA** library could serve as the entry point since they possessed a high quantity of BC zwitterion-binding sequences. The selection could be performed at physiological temperature (37 ˚C) in biological buffer without additional supplements to improve the target aptamer's thermostability and cation dependence. During the selection, progressively lowering the dye concentration and applying a stricter sorting gate based on the NIR fluorescence could be performed to ensure the successful isolation of high-affinity binders with improved brightness. By screening and characterization of the highly enriched sequences, we anticipate an improved version of the color-shifting aptamer-fluorophore module that is capable of more complex biological applications

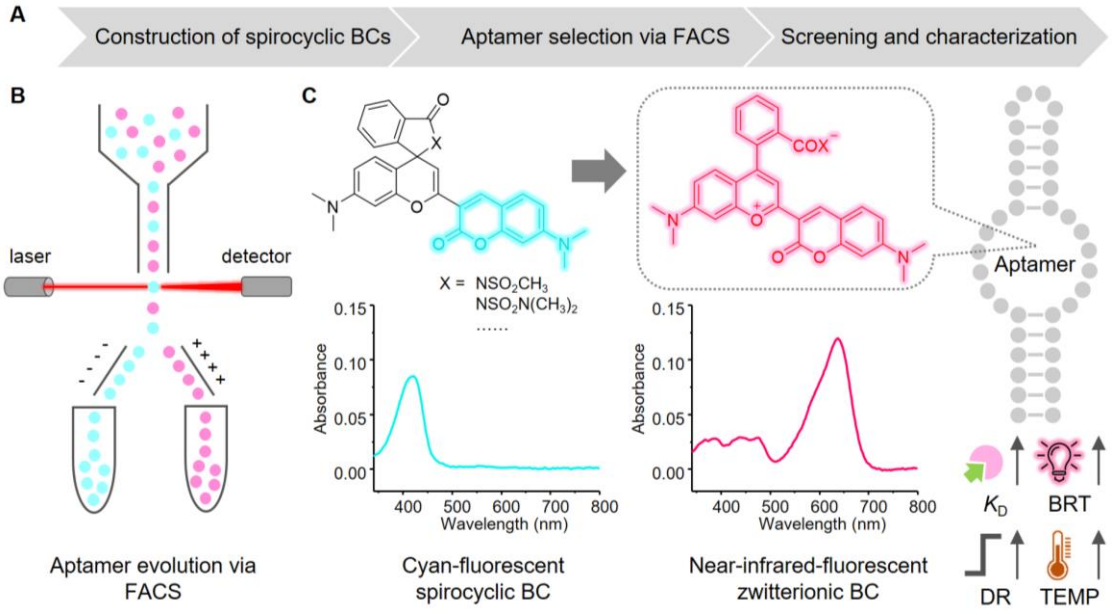


Figure 75. Evolution of the next-generation of color-shifting aptamer-fluorophore modules through FACS. (A) Workflow for the generation of a new color-shifting aptamer-fluorophore module. (B) Aptamer selection using FACS. The highly NIR-fluorescent cells will be collected and used for the next round of selection. (C) Illustration of a new color-shifting aptamer-fluorophore module with improved biophysical properties. The representative cyan-fluorescent spirocyclic BCs are shown. The aptamer specifically binds to their zwitterionic forms, shifting the fluorescence from cyan to NIR. K_D, BRT, DR, TEMP represent dissociation constant, brightness, dynamic range, and thermal stability, respectively.

4.2.2 Color-shifting aptamer-fluorophore module for RNA sensor development

Genetically encoded protein sensors have been widely used for real-time imaging and monitoring of small molecules in living cells[142-145]. Nevertheless, the design and development of a sensing domain that undergoes a conformational change upon target binding are complicated and time-consuming[125]. Recently, RNA-based biosensors comprising a target-binding aptamer and a FLAP have been developed to detect various small molecules and proteins[125, 146-147]. However, precise determination and monitoring of target molecules' intracellular concentrations are difficult with these single-color fluorogenic RNA sensors. Our color-shifting aptamer-fluorophore module, as a ratiometric fluorescence reporter, could alleviate the signal fluctuations derived from variations in cellular uptake of the probe, cell morphology, and the environmental conditions, thereby offering more accuracy.

Although a tetracycline sensor was successfully developed *in vitro* (chapter 3.1.9), better variants that enable a larger dynamic range and an improved EC_{50} should be generated for practical use in cells (Figure 76A). By varying the transducer stem's length and stability, as demonstrated in chapter 3.1.9.3, more tetracycline sensors could be designed and screened for emission ratio changes in response to tetracycline. Moreover, a doped library based on the best sensor candidate sequence could be generated and used for sensor evolution through FACS. The cells demonstrating the largest dynamic range to tetracycline would be collected and analyzed. Besides, the design strategy of ratiometric RNA sensors could be expanded to other interesting targets, including cellular metabolites, signaling molecules, and synthetic compounds that are difficult to detect by conventional small-molecule or protein sensors. The cognate RNA aptamers as recognition module could be obtained either from reported literature or via SELEX.

Once the aptamer sensor platform is established, its functionality in living mammalian cells would be examined. To overcome the limitations such as low expression and rapid degradation of an RNA sensor in mammalian cells, a twister-optimized RNA for durable overexpression (Tornado) system could be exploited (Figure 76B)[127]. The designed autocatalytic RNA containing the sensor sequence could be processed to circular sensor RNAs with high stability and expression level. This ratiometric RNA senor platform could serve as a valuable addition to the intracellular sensing tools.

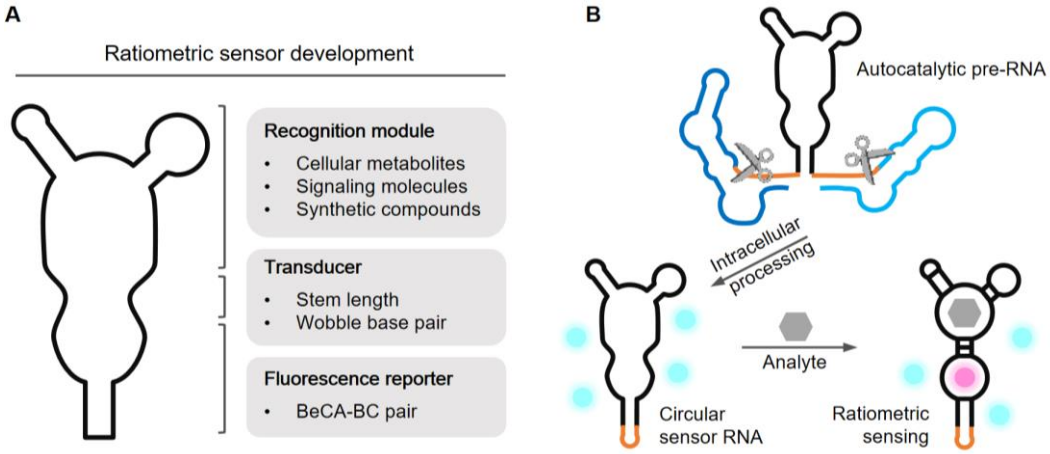


Figure 76. Ratiometric sensor development for intracellular sensing. (A) Development of a ratiometric sensor using the color-shifting aptamer-fluorophore module. The recognition module could be changed for detecting different targets. The transducer module could be modified by varying the stem length or stability to fine-tune the sensor's dynamic range. (B) Circular sensor RNA for intracellular sensing. The pre-RNA could be self-cleaved and ligated to form the circular RNA that induces ratiometric signal upon target binding with exceptionally high concentration inside cells.

4.2.3 Analysis, truncation, characterization of SiR-binding aptamers

We identified several new RNA sequences, which showed improved biophysical properties compared to SiRA against the ligand SiR-PEG$_3$-NH$_2$. Nevertheless, further analysis and characterization of these sequences need to be carried out to generate the next generation of SiR-binding NIR FLAPs.

In addition to L1 and L2, NGS or larger-scale Sanger sequencing of the DNA pools generated from the doped library (i.e., L3) should be conducted. "Doping" has been proven to be an efficient way to generate better aptamer variants[43, 59, 139]. By analyzing the sequencing results of L3, new patterns that promote tighter and stronger interaction between the aptamer and the ligand could be foreseen. Meanwhile, the truncation of the L2-1 and L2-8 could be continued since a smaller imaging tag is always preferred. Moreover, the best aptamer-fluorophore pair should be carefully characterized with respect to brightness (quantum yield × molar extinction coefficient), cation and temperature dependence, photostability, etc.

4.2.4 Spirocyclization-based fluorogenic dyes for FLAPs

In the last decade, a new class of fluorogenic dyes based on the intramolecular spirocyclization equilibrium was developed and successfully applied in live-cell imaging[73, 81-82]. The spectral property and the fluorogenicity of these rhodamine derivatives could be fine-tuned by modifying either the chromophore or the intramolecular nucleophile (chapter 1.3.1.2 and chapter 1.3.3)[82, 93]. These highly fluorogenic dyes with emission spanning the whole visible spectral range are particularly interesting candidates for RNA imaging since they possess low background signal and aptamer-induced fluorescence with a high dynamic

range. These fluorophores could be screened as substrates for a particular aptamer to generate a promiscuous FLAP or used as individual ligands for generating a palette of orthogonal FLAPs.

Substrate promiscuity

FLAPs comprised of a promiscuous aptamer and a series of binding ligands spanning the whole visible spectral range are beneficial since the aptamer's color could be interchanged by simply supplementing the respective ligand[50]. The use of tetramethylrhodamine and carbopyronine demonstrated the substrate promiscuity of SiRA as they both showed similar affinity towards SiRA relative to SiR[65]. The substitutions of oxygen and carbon atoms to the central silicon atom shifted the emission colors of the aptamer-fluorophore complex. Nevertheless, the binding affinity was not remarkably disturbed due to the minor structural variation on the fluorophore. Analogously, the replacement of the carboxyl group to electron-deficient amides increases the fluorogenicity of these rhodamine-based fluorophores. Yet, they still hold the potential to preserve the intact binding affinity as demonstrated by **BeCA-BC6** (evolved from **BeCA-BC1**; chapter 3.1). Therefore, we believe that by screening these fluorogenic rhodamine derivatives towards the newly evolved SiR-binding aptamer, a promiscuous FLAP with a broader spectral range would be achieved (Figure 77)[50].

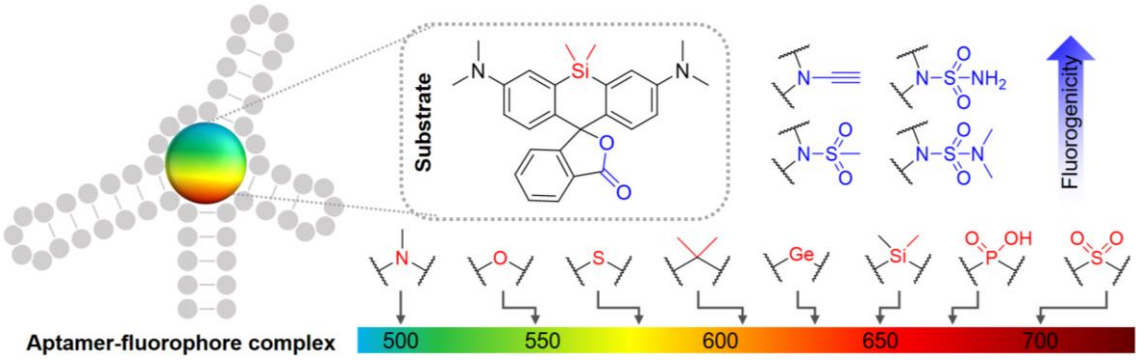


Figure 77. Expanding the spectral range of a FLAP. The structure of SiR is shown. The emission of the aptamer-fluorophore complex could be tuned by using different rhodamine derivatives (indicated in red). The fluorogenicity of the aptamer-fluorophore complex could be achieved by modifying the substrate's intramolecular nucleophile (indicated in blue).

Orthogonal labeling

Multicolor labeling of RNAs with orthogonal FLAPs would facilitate the simultaneous visualization of various RNA targets. Although several FLAPs have been developed, their intracellular performances are still limited by the unsatisfying photophysical properties. The FACS-based aptamer selection platform developed in chapter 3.2 offered an excellent tool to evolve light-up aptamers for those spirocyclization-based fluorophores with superior biophysical characteristics. The aptamer's specificity towards its cognate

substrate could be improved by applying counter selections using the substate's analogs. Thereby, we would obtain a palette of bright FLAPs with high sensitivity and specificity for live-cell imaging.

4.2.5 A dimerization-based aptamer-fluorophore module

The dimerization of SRB2 and its cognate ligand sulforhodamine B generated o-Coral-Gemini-561, displaying an enhanced fluorescence performance *in vitro* and in cells[60]. The newly evolved SiR-binding aptamers through FACS, although possessing an improved binding affinity compared to SiRA, could be further optimized by a dimerization strategy (Figure 78). A SiR-dimer formed by conjugation of two SiRs via a flexible linker could reduce the background noise from the free probe due to the aggregation-induced self-quenching. More importantly, by exploiting an aptamer dimer, the binding of one SiR out of the SiR-dimer to the cognate aptamer could facilitate the other binding event. An affinity down to the low-nanomolar/picomolar range could be expected due to the avidity effect. This dimeric system with an ultra-high binding affinity and an increased fluorogenicity would bring the aptamer-fluorophore's imaging performance to the next level.

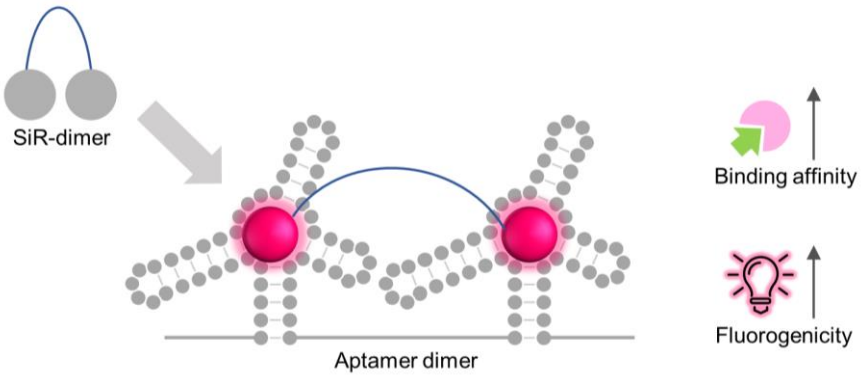


Figure 78. A dimeric aptamer-fluorophore module with an increased binding affinity and fluorescence enhancement.

4.2.6 Endogenous RNA imaging in mammalian cells

Real-time imaging of endogenous RNAs could help understand the intricate RNA dynamics inside cells[19, 32, 148]. In this thesis, we applied the **BeCA-BC6** system to visualize target RNAs in bacterial cells where **BeCA**-tagged RNAs were expressed under the control of a non-native promoter (T7) within a pET expression vector (Figure 79). SiRA also demonstrated its imaging capability using the same expression system in *E. coli* cells[65]. Nevertheless, labeling of endogenous RNAs without artifacts coming from an over-expression system is preferred to precisely understand the localization and dynamics of RNAs of interest in living cells. This is even more challenging in mammalian cells due to the low-copy number of most endogenous RNAs and a more complex

intracellular environment[149]. Once the newly evolved aptamer-fluorophores are obtained (for both BC- and SiR-binding FLAPs), we could start testing their performances in mammalian cells. 5S ribosomal RNA is reported to possess a relatively long half-live and high abundance, therefore a good model transcript for assessing a novel FLAP in mammalian cells at the early stage[41, 58, 149]. First, the expression of aptamer-tagged RNA in mammalian cells could be controlled under a native promoter with a relatively high copy number using a transfection system[41]. If the imaging results are promising, we will move on to label endogenous 5S ribosomal RNA with the help of genome editing tools, such as CRISPR-Cas9. The endogenous RNAs could be labeled by the knock-in of the aptamer sequence to their respective native genomic loci. With the development of RNA imaging tools and imaging techniques such as super-resolution microscopy, we believe that a detection limit down to the single-molecule level could be achieved using the super-bright, photostable, high-affinity aptamer tandem fused to any RNA of interest.

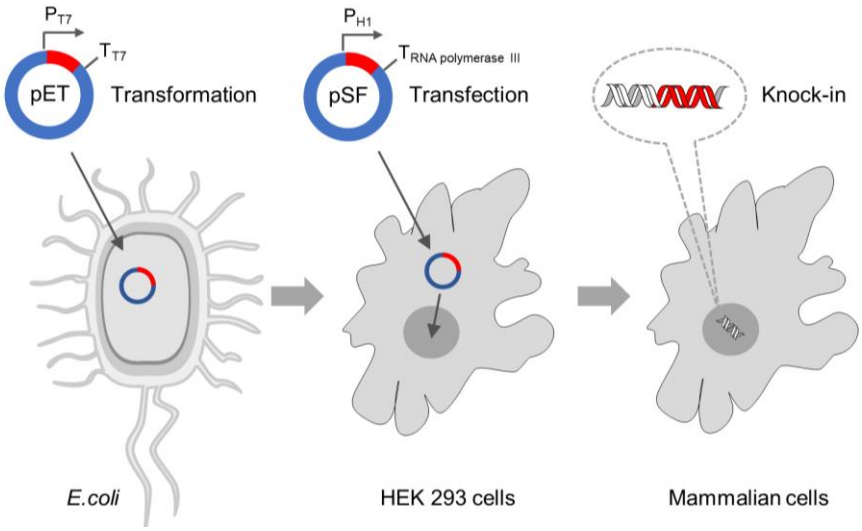


Figure 79. RNA imaging in living cells. The aptamer-tagged RNAs (gene indicated in red) are currently transformed and imaged in bacterial cells but will be transfected and imaged in mammalian cells. The final goal is to knock-in the aptamer tag to the gene of interest and image the respective endogenous RNA in living cells.

5 Materials and methods

The majority of the content described in this chapter was covered in the electronic supplementary of an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.

5.1 General materials and standard methods

5.1.1 Reagents

5.1.1.1 Chemical reagents

All chemical reagents and solvents for synthesis were purchased from Sigma Aldrich, Alfa Aesar, and Tokyo Chemical Industry and used without further purification. Deuterated solvents for NMR spectroscopy were purchased from Euriso-Top.

5.1.1.2 Biology reagents

All chemicals (molecular biology grade) for biological experiments were purchased from Sigma-Aldrich and Carl Roth unless noted otherwise. The essential reagents, enzymes, kits, plasmids, and cells are listed in Table 11.

Category	Name	Supplier
Chemicals	Ethidium bromide (EtBr)	Carl Roth
	Agarose	Carl Roth
	Ammonium persulfate (APS)	Sigma Aldrich
	Tetramethylethylenediamine (TEMED)	Sigma Aldrich
	ROTI Aqua-P/C/I (for RNA)	Carl Roth
	phenol/chloroform/isoamyl 25:24:1 (for DNA)	Sigma Aldrich
	dNTP mix (10 mM each)	Thermo Fisher Scientific
	NTPs (ATP, GTP, CTP, UTP)	Carl Roth
	Dithiothreitol (DTT)	Sigma Aldrich
	Formamide	Sigma Aldrich
	Ethanolamine	Sigma Aldrich
	Ethylenediaminetetraacetic acid (EDTA)	Sigma Aldrich
	Triton X-100	Merck Millipore
	Tween 20	Sigma Aldrich
	Spermidine	Sigma Aldrich
	Bromophenol blue	Waldeck
	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Sigma Aldrich

	Trizma base	Sigma Aldrich
	Boric acid	NeoFROXX
	Synthetic DNA oligos	Integrated DNA Technologies
	NHS-activated sepharose 4 fast flow	GE Healthcare Life Sciences
	Streptavidin immobilized on agarose	Sigma Aldrich
	Glycogen	Thermo Fisher Scientific
	Ampicillin sodium salt	Carl Roth
	Kanamycin sulfate (from *Streptomyces kanamyceticus*)	Sigma Aldrich
	Isopropyl-β-D-thiogalactopyranoside (IPTG)	Sigma Aldrich
	d-Glucose	Sigma Aldrich
	Poly-D-lysine	Sigma Aldrich
	Lysogeny broth (LB)	Lennox, Carl Roth
	LB broth with agar	Lennox, Carl Roth
	M9 broth	Fluka
	RNAzol RT	Sigma Aldrich
Enzymes	*Taq* DNA polymerase	Lab-prepared
	T7 RNA polymerase	Lab-prepared
	DNase I, RNase-free	Thermo Fisher Scientific
	SuperScript IV reverse transcriptase	Thermo Fisher Scientific
	Restriction enzymes (AatII, SalI, XhoI)	Thermo Fisher Scientific
	Q5 high-fidelity DNA polymerase	New England Biolabs
	FastAP thermosensitive alkaline phosphate	Thermo Fisher Scientific
	T4 DNA ligase	Thermo Fisher Scientific
	T4 polynucleotide kinase	Thermo Fisher Scientific
Kits	QIAquick PCR purification kit	Qiagen
	Fast DNA End Repair Kit	Thermo Fisher Scientific
	GeneJET plasmid miniprep kit	Thermo Fisher Scientific
Plasmids	pDisplay-AP-CFP-TM	Addgene
	pET-28a(+)	Merck KGaA
	pET His6 GFP TEV LIC cloning vector (1GFP)	Addgene
	pTAC-MAT-Tag-2	Thermo Fisher Scientific
Cells	BL21 Star (DE3) competent *E. coli*	Thermo Fisher Scientific
	DH5α competent *E. coli*	Thermo Fisher Scientific

Table 11. List of biology reagents used in this thesis.

5.1.1.3 Buffers

The important lab-prepared buffers used for biological experiments are listed in Table 12.

10 × standard *Taq* reaction buffer	200 mM Tris-HCl, 500 mM KCl, pH 8.4
10 × *In vitro* transcription (IVT) buffer	400 mM Tris-HCl, 10 mM spermidine, 220 mM $MgCl_2$, 0.1% Triton X-100, pH 8.4
6 × Aptamer selection buffer (ASB)	120 mM HEPES, 30 mM $MgCl_2$, 750 mM KCl, pH 7.4
10 × Tris/Borate/EDTA (TBE) buffer	1 M Tris-borate, 20 mM EDTA, pH 8.3
2 × RNA gel loading dye	1 × TBE, 90% (v/v) formamide, 0.1% (w/v) bromophenol blue
6 × DNA gel loading dye	30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF

Table 12. List of buffers used in this thesis.

5.1.2 Instruments and materials

The essential materials and instruments are listed in Table 13.

Category	**Name**	**Supplier**
Materials	Polygram Sil G/UV254 TLC plates	Macherey-Nagel
	Silica gel (pore size: 60 Å; particle size: 40 – 63 µm)	Sigma Aldrich
	C18 Phenomenex Luna (5 µm)	Phenomenex
	Ultra-Micro cells 105.251-QS	Hellma
	Ultra-Micro Cell 105.202-QS	Hellma
	µ-Slide 8-well glass chamber	ibidi
Instruments	Analytical balances	Mettler Toledo
	pH meter	Mettler Toledo
	1100 HPLC	Agilent Technologies
	Alpha 1-4 LD lyophilizer	Martin Christ Gefriertrocknungsanlagen
	Vacuum pump	Vacuubrand
	Varian Mercury Plus 300/500 MHz	Varian
	Bruker microTOF-QII mass spectrometer	Bruker
	Milli-Q Synthesis A10	Merck Millipore
	Cary 50 UV-visible spectrophotometer	Agilent Technologies
	FP-6500 fluorescence spectrometer	JASCO
	NanoDrop ND-1000 spectrophotometer	Thermo Fisher Scientific
	T100 thermal cycler	Bio-Rad
	ThermoMixer	Eppendorf
	Centrifuge 5804 R	Eppendorf

	MIKRO 200 centrifuge	Hettich
	AlphaImager TM 2200	Alpha Innotech
	MaxQ 8000 Orbital Shaker	Thermo Fisher Scientific
	A1R confocal microscope	Nikon
	Typhoon 9400	GE Healthcare Life Sciences
	MicroPulser Electroporator	Bio-Rad
	FACS Aria III	Becton Dickinson

Table 13. List of materials and instruments used in this thesis.

5.1.3 Software and webserver

The important software and webserver are listed in Table 14.

Category	**Name**	**Supplier/Website**
Software	ChemDraw Professional 16.0	PerkinElmer
	OriginPro 2015	OriginLab
	MestReNova-9.0.1	Mestrelab Research
	GraphPad Prism 6	GraphPad Software
	ImageJ/Fiji	National Institutes of Health
	VARNA v3-93	Developed by Yann Ponty
Webserver	MultAlin	http://multalin.toulouse.inra.fr/multalin/
	RNAstructure	https://rna.urmc.rochester.edu/RNAstructureWeb/

Table 14. List of software and webserver used in this thesis.

5.1.4 PCR amplification

The PCR amplifications in this thesis were performed using the standard protocol described below and purified via QIAquick PCR purification kit unless otherwise noted (Table 15 and Table 16).

Component	**50 µL reaction**	**Final concentration**
Template DNA	variable	variable (pg - ng)
Forward primer (100 µM)	1 µL	2 µM
Reverse primer (100 µM)	1 µL	2 µM
dNTPs (10 mM)	2.5 µL	0.5 mM
$MgCl_2$ (50 mM)	4 µL	4 mM
10 × standard *Taq* reaction buffer	5 µL	1 ×
Taq DNA polymerase (lab-prepared)	1 µL	2% (v/v)
Nuclease-free water	to 50 µL	

Table 15. Reaction setup for a standard PCR.

Step	Temperature	Time	Repeat
Initial denaturation	95 ˚C	2 min	1
Denaturation	95 ˚C	30 sec	6 - 36 cycles
Annealing	50-72 ˚C	30 sec	
Elongation	72 ˚C	30 sec	
Final extension	72 ˚C	5 min	1
Hold	4 ˚C	∞	1

Table 16. Thermocycling conditions for a routine PCR.

5.1.5 *In vitro* transcription

The IVTs in this thesis were performed according to the standard protocol described below unless noted otherwise (Table 17). Reaction mixtures with varying scales were incubated at 37 ˚C for 4 h, followed by DNase I treatments (1 U per 1 µg template DNA) for 30 min. RNAs were purified by denaturing polyacrylamide gel electrophoresis (PAGE) unless noted otherwise.

Component	250 µL reaction	Final concentration
DNA template	variable	variable (ng - µg)
ATP (100 mM)	10 µL	4 mM
GTP (100 mM)	10 µL	4 mM
CTP (100 mM)	10 µL	4 mM
UTP (100 mM)	10 µL	4 mM
DTT (1 M)	2.5 µL	10 mM
10 × IVT buffer	25 µL	1 ×
T7 RNA polymerase	5 µL	2% (v/v)
Nuclease-free water	to 250 µL	

Table 17. Reaction setup for a standard IVT (preparative scale).

5.1.6 Phenol/chloroform/isoamyl alcohol purification

Large-scale PCR and IVT products were purified by phenol/chloroform/isoamyl alcohol (PCI) extraction (chapter 5.3.1). One volume of the PCI solution (25:24:1, pH 8 for DNA or pH 4.5-5 for RNA, respectively) was added to the sample. After mixing thoroughly, the suspension was centrifuged at 9000 rpm for 2 min to allow phase separation. The aqueous phase was removed, placed into a new tube, and washed twice with one volume of chloroform. The final aqueous solution was supplemented with NaAc (final concentration; 0.3 M, pH 5.5) and 2 volumes of absolute ethanol (or 1 volume of

isopropanol) for DNA or 2.5-3 volumes of absolute ethanol (or 1 volume of isopropanol) for RNA. The mixture was incubated at − 80 ˚C for at least 1 h or − 20 ˚C overnight to allow precipitation. The solution was then centrifuged (11000 rpm) at 4 ˚C for 1 h in a high-speed centrifuge. After decanting, the DNA/RNA pellet was rinsed twice with 75% ethanol, dried in the air, and redissolved in water for further use.

5.1.7 Gel electrophoresis

5.1.7.1 Agarose gel electrophoresis

PCR products were checked by agarose gel electrophoresis unless noted otherwise. The agarose gel was prepared by dissolving 1-2% (w/v) agarose (depending on the DNA size) in 1 × TBE buffer under microwave heating. The gel was cooled down to 50 ˚C before adding EtBr to a final concentration of 1 µg/mL. The DNA sample was examined in the solidified gel (120 V for 45 min) and visualized in AlphaImager under UV light.

5.1.7.2 Denaturing polyacrylamide gel electrophoresis

RNA purifications were performed by denaturing polyacrylamide gel electrophoresis (PAGE) unless noted otherwise. The mixture containing 10% of urea-acrylamide, 0.1% (w/v) APS and 0.1% (v/v) TEMED was prepared for gel polymerization. The RNA sample was then loaded and examined in the polyacrylamide gel at 25 W for 2 h in 1 × TBE buffer. The RNA band was visualized by UV shadowing and excised. The target RNA was extracted from the gel using the "crush and soak" method in a NaAc (0.3 M pH = 5.5) solution. After precipitation with 3 volumes of absolute ethanol, the eluted RNA was resuspended in water for further use. For analysis purposes, instead of RNA elution, the gel was stained with EtBr for 30 min (5 µL in 100 mL TBE buffer) and imaged in Typhoon (laser: 532 nm; band-pass (BP) filter: 610 ± 15 nm).

5.1.8 RNA folding

The desired amount of RNA was diluted in ultrapure water, incubated at 75 °C for 2 min, and slowly cooled down to 25 °C within 10 min. 1/5 volume of 6 × ASB was added to the solution, and the mixture was kept at 25 °C for 15 min.

5.1.9 Transformation

The bacterial transformations in this thesis were mostly conducted using the chemical-based approach. The only exceptions were described in chapter 3.2.2.4, where the bacterial libraries for FACS were prepared by electroporation.

5.1.9.1 Chemical transformation

The chemically competent cells (DH5α or BL21 Star [DE3]) were thawed on ice for 15 min. 1-2 μL of plasmid DNA (usually 10 pg - 100 ng) was added to the competent cells (40 μL) in a microcentrifuge tube. The cell-DNA mixture was incubated on ice for 30 min, followed by a heat-shock at 42 ˚C for 45 sec, and then placed on ice for 2 min. After addition of 250 μL of LB medium (without antibiotic), the bacteria were gown at 37 °C for 45 min in a shaking incubator. A small portion (50 μL) of the transformants were placed on an LB agar plate containing the appropriate antibiotic (30 μg/mL kanamycin or 100 μg/mL ampicillin).

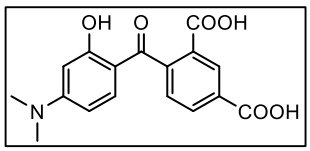
5.1.9.2 Electroporation

Freshly prepared or thawed BL21 Star (DE3) electrocompetent cells (40 μL) were mixed with 1-2 μL of DNA in a pre-cold microcentrifuge tube on ice. After incubating for 1 min, the cell-DNA mixture was transferred to the bottom of a pre-cold electroporation cuvette (gap size: 0.1 cm). The cuvette was then placed on the chamber slide of the electroporator and pulsed once. LB medium (1 mL without antibiotic) was immediately added to the bacteria mixture. After gentle resuspension, the cells were transferred to a 15 mL polypropylene tube and incubated at 37 °C for 1 h with shaking (170 rpm). Finally, the cells were placed into a selective medium (containing 30 μg/mL kanamycin or 100 μg/mL ampicillin) for further use.

5.2 Synthesis and characterization

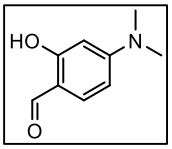
The synthetic procedures described in this chapter were covered in the electronic supplementary of an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.
All the reactions were monitored by thin-layer chromatography (TLC) performed on Polygram Sil G/UV254 TLC plates. Flash column chromatography was conducted using silica gel with a pore size of 60 Å and a particle size range of 40 – 63 μm. All NMR spectra were recorded on a Varian Mercury Plus 300 or 500 MHz. The chemical shifts are reported in parts per million (ppm), and the coupling constants are given in hertz (Hz). High-resolution mass spectrometry (HRMS) was performed on a Bruker microTOF-QII mass spectrometer.

5.2.1 Synthesis of BC1-NH_2 and BC1-biotin

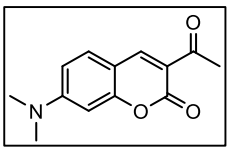


4-(4-(dimethylamino)-2-hydroxybenzoyl)isophthalic acid (1): 3-Dimethylaminophenol (2.20 g, 16.0 mmol) was dissolved in toluene (50 mL) and heated to 60 °C. Powdered trimellitic anhydride (3.84 g, 20.0 mmol) was added while stirring. The mixture was refluxed for 24 h and cooled down to room

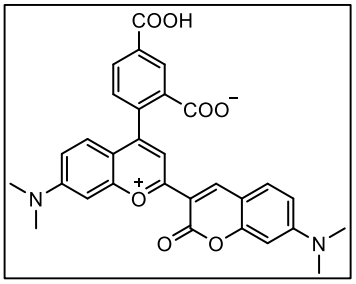
temperature. The residue was filtered off, washed with toluene (3 × 30 mL), dissolved in methanol (80 mL), and refluxed for 10 min. Then, acetic acid (20 mL) was added, and the solvent was removed under reduced pressure at 40 °C. The solid was refluxed in methanol (50 mL) for 2 h and kept at 4 °C overnight. The crystals formed were filtered off and washed with ice-cold methanol (50 mL) to give pure compound **1** as brown crystalline solid.
Yield: 724 mg, 2.24 mmol, 14%.
^{1}H NMR (300 MHz, C_2D_6OS) δ 8.45 (s, 1H), 8.16 (d, J = 7.8 Hz, 1H), 7.49 (d, J = 7.9 Hz, 1H), 6.77 (d, J = 9.1 Hz, 1H), 6.17 (d, *J* = 7.4 Hz, 1H), 6.08 (s, 1H), 2.97 (s, 6H).
^{13}C NMR (75 MHz, C_2D_6OS) δ 198.0, 166.5, 164.6, 156.2, 144.2, 134.2, 133.1, 132.1, 131.1, 130.4, 128.7, 109.9, 104.9, 97.4.
HRMS (ESI): m/z calc. for $C_{17}H_{15}NNaO_6$ 325.0792; found 325.0796, $[M+Na]^+$.



4-(dimethylamino)-2-hydroxybenzaldehyde (2): Phosphorus(V) oxychloride (4.67 mL, 50.0 mmol) was added dropwise into dimethylformamide (8 mL) at 0 °C under argon. The mixture was stirred for 5 min at 0 °C and then 40 minutes at room temperature. A solution of 3-(dimethylamino)phenol (3.43 g, 25.0 mmol) in dimethylformamide (2 mL) was added dropwise. The mixture was heated and stirred at 70 °C for 1 h. The solution was poured onto crushed ice (approx. 50 g) and extracted with ethyl acetate. The combined organic phases were dried over anhydrous magnesium sulfate and evaporated under reduced pressure. The crude product was purified by column chromatography using hexane:ethyl acetate mixtures (10:1 to 10:4).
Yield: 2.80 g, 17.0 mmol, 68%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 9.52 (s, 1H), 7.26 (d, *J* = 5.7 Hz, 1H), 6.27 (dd, *J* = 8.9, 2.4 Hz, 1H), 6.07 (d, *J* = 2.4 Hz, 1H), 3.06 (s, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 192.4, 164.0, 156.1, 135.1, 111.7, 104.5, 97.2, 40.1.
HRMS (ESI): m/z calc. for $C_9H_{11}NNaO_2$ 188.0682; found 188.0684, $[M+Na]^+$.



3-acetyl-7-(dimethylamino)chromen-2-one (3): Compound **2** (407 mg, 2.50 mmol) and ethyl acetoacetate (403 mg, 3.10 mmol) were dissolved in absolute ethanol (5 mL). Piperidine (9 drops) was slowly added, and the solution was heated under reflux for 5 h. The mixture was cooled at 4 °C overnight, leading to the formation of a yellow precipitate, which was filtered off and washed with ice-cold ethanol to provide the pure product.
Yield: 450 mg, 1.95 mmol, 78%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.44 (d, J = 0.7 Hz, 1H), 7.41 (d, J = 8.9 Hz, 1H), 6.63 (dd, J = 8.9, 2.5 Hz, 1H), 6.46 (d, J = 2.4 Hz, 1H), 3.12 (s, 7H), 2.67 (s, 3H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 195.7, 160.7, 158.3, 154.9, 148.0, 131.6, 116.8, 110.0, 108.4, 97.0, 40.2, 30.6.
HRMS (ESI): m/z calc. for $C_3H_{13}NNaO_3$ 254.0788; found 254.0778, $[M+Na]^+$.



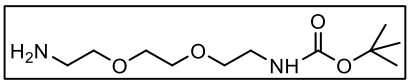
Compound 4: Compounds **1** (659 mg, 2.00 mmol) and **3** (462 mg, 2.00 mmol) were dissolved in concentrated sulfuric acid (5 mL) and stirred for 6 h at 90 °C. The mixture was cooled down to room temperature, and ice (approx. 15 g) was added carefully. After adding 70% perchloric acid (1 mL), the mixture was filtered, and the precipitate was washed with water. The crude product was air-dried overnight and further purified by column chromatography using dichloromethane:methanol mixtures (10:1 to 10:3).

Yield: 356 mg, 680 µmol, 34%.

^{1}H NMR (300 MHz, C_2D_6OS) δ 9.11 (s, 1H), 8.64 (d, J = 1.7 Hz, 1H), 8.32 (dd, J = 8.0, 1.8 Hz, 1H), 8.11 (s, 1H), 7.66 (dd, J = 15.7, 8.5 Hz, 2H), 7.23 (s, 3H), 6.95 (d, *J* = 8.8 Hz, 1H), 6.68 (d, J = 2.3 Hz, 1H), 3.28 (s, 6H), 3.17 (s, 6H).

^{13}C NMR (126 MHz, C_2D_6OS) δ 166.4, 162.0, 158.4, 158.1, 158.1, 157.2, 156.9, 146.7, 140.1, 133.5, 133.1, 133.0, 131.9, 131.3, 130.7, 129.4, 118.0, 116.0, 112.5, 111.7, 110.2, 106.0, 97.3, 96.8, 40.7, 40.2.

HRMS (ESI): m/z calc. for $C_{30}H_{25}N_2O_7$ 525.1656; found 525.1639, $[M+H]^+$.



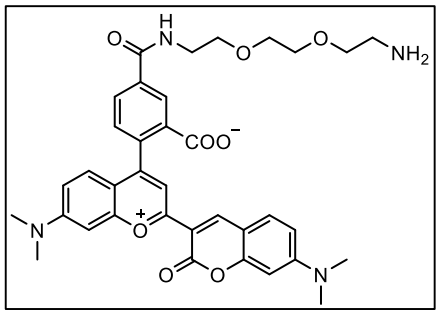
N-Boc-2,2′-(ethylenedioxy)diethylamine (5): To a solution of 2,2′-(Ethylenedioxy)bis(ethylamine) (4.45 g, 30.0 mmol) in anhydrous dichloromethane (200 mL) at 0 °C was added di-tert-butyl dicarbonate (1.20 g, 5.50 mmol) dissolved in anhydrous dichloromethane (100 mL) dropwise over a period of 1 h. The reaction mixture was allowed to warm up to room temperature and stirred for 2 h. The solution was washed with brine (2 × 100 mL), dried over anhydrous magnesium sulfate, and the solvent was removed under reduced pressure. The transparent and highly viscous liquid was lyophilized and used without further purification.

Yield: 967 mg, 3.90 mmol, 78%.

^{1}H NMR (300 MHz, $CDCl_3$) δ 3.61 (s, 4H), 3.52 (dt, *J* = 9.4, 5.2 Hz, 4H), 3.31 (q, *J* = 5.4 Hz, 2H), 2.87 (t, J = 5.2 Hz, 2H), 1.43 (s, 9H).

^{13}C NMR (75 MHz, $CDCl_3$) δ 156.0, 79.1, 73.5, 70.2, 41.8, 40.3, 28.4.

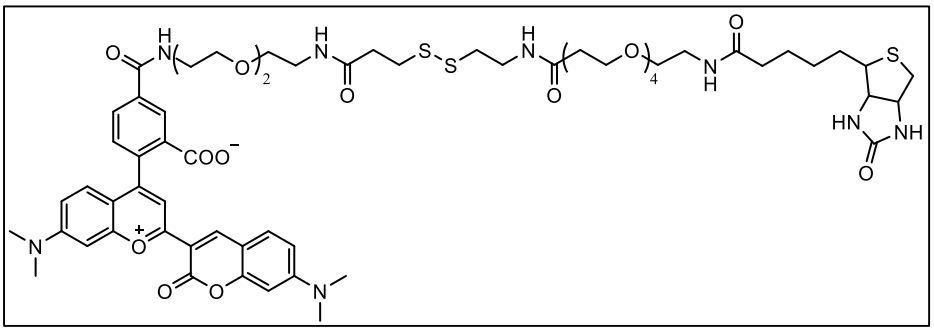
HRMS (ESI): m/z calc. for $C_{11}H_{25}N_2O_4$ 249.1809; found 249.1817, $[M+H]^+$.



BC1-NH$_2$ (6): Compound **4** (105 mg, 200 µmol) and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (77.0 mg, 240 µmol) was dissolved in anhydrous dimethylformamide (1 mL). After 30 min stirring at room temperature, N,N-diisopropylethylamine (77.6 mg, 600 µmol), and **5** (99.2 mg, 400 µmol) was added. The solution was stirred overnight. The next day, the mixture was diluted with water (100 mL) and extracted with ethyl acetate (5 × 100 mL). The combined organic phases were

dried over anhydrous magnesium sulfate and evaporated under reduced pressure. The obtained residue was purified by column chromatography using dichloromethane/methanol (10:0 to 10:2). The intermediate was dissolved in trifluoroacetic acid:dichloromethane mixture (1:4) and stirred at room temperature for 2 h under argon. The solvent was removed under vacuum, and the blue residue was purified by RP-HPLC.
Yield: 16.0 mg, 24.5 µmol, 12%.
^{1}H NMR (500 MHz, C_2D_6OS) δ 8.68 (s, 1H), 8.39 (d, J = 1.5 Hz, 1H), 8.23 (dd, J = 8.0, 1.6 Hz, 1H), 7.69 (d, J = 8.9 Hz, 1H), 7.44 (d, J = 8.1 Hz, 1H), 6.85 (dd, J = 9.0, 2.5 Hz, 1H), 6.69 (d, J = 1.5 Hz, 1H), 6.61 (d, J = 2.4 Hz, 1H), 6.54 (d, J = 1.4 Hz, 2H), 6.38 (s, 1H), 3.61 – 3.54 (m, 8H), 3.47 (q, J = 5.9 Hz, 4H), 3.08 (s, 6H), 2.97 (s, 6H).
^{13}C NMR (126 MHz, C_2D_6OS) δ 168.7, 165.3, 158.9, 156.1, 155.9, 154.3, 152.3, 152.2, 147.4, 142.0, 136.5, 134.8, 130.8, 128.4, 126.4, 124.6, 123.7, 110.7, 110.4, 110.3, 108.4, 105.2, 99.6, 98.7, 97.0, 84.1, 70.1, 69.9, 69.2, 68.3, 40.6, 40.4, 40.2, 40.1.
HRMS (ESI): m/z calc. for $C_{36}H_{39}N_4O_8$ 655.2762; found 655.2801, $[M+H]^+$.

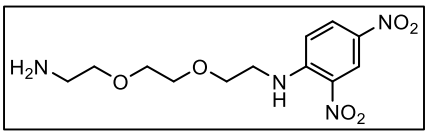
BC1- Biotin (7): To a solution of **BC1-NH$_2$** (1.30 mg, 2.00 µmol) in anhydrous



dimethylformamide (100 µL), NHS-SS-PEG$_4$-biotin (1.80 mg, 2.40 µmol) was added. After stirring for 20 min, triethylamine (364 µg, 3.60 µmol) was added, and the mixture was stirred for another 20 min. The product was purified via RP-HPLC.
Yield: 246 µg, 0.19 µmol, 10%.
HRMS (ESI): m/z calc. for $C_{62}H_{82}N_8NaO_{16}S_3$ 1313.4903; found 1313.4946, $[M+Na]^+$.

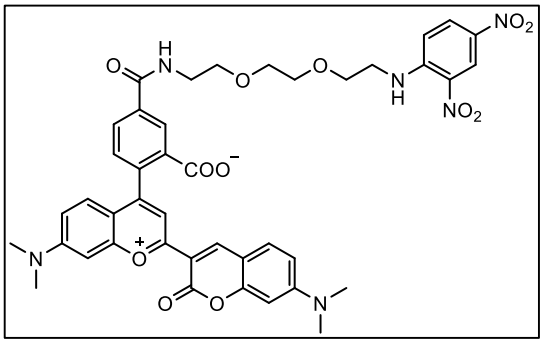
5.2.2 Synthesis of BC1-DN

N-(2-(2-(2-Aminoethoxy)ethoxy)ethyl)-2,4-dinitroaniline (8): 1-Fluoro-2,4-dinitrobenzene (205 mg, 1.10 mmol) dissolved in dichloromethane (5 mL) was added dropwise to



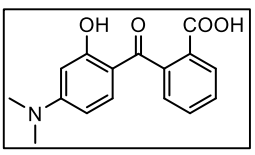
a solution of 2,2′-(Ethylenedioxy)bisethylenediamine (1.48 g, 10.0 mmol) in dichloromethane (10 mL) at 0 °C. The mixture was stirred for 30 minutes at room temperature after the addition was complete and then diluted with a mixture of dichloromethane (100 mL) and water (100 mL). The organic phase was separated and mixed thoroughly with 0.1 M hydrochloric acid (50 mL). The aqueous phase was recovered, and the pH was adjusted to 12 with sodium hydroxide. The

product was extracted into the organic phase again using dichloromethane (3 × 50 mL). The organic phases were combined, washed with brine (100 mL), dried over anhydrous magnesium sulfate, and evaporated to yield pure compound **8**.
Yield: 310 mg, 0.98 mmol, 92%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 9.12 (d, J = 2.7 Hz, 1H), 8.25 (dd, J = 9.5, 2.7 Hz, 1H), 6.93 (d, J = 9.5 Hz, 1H), 3.83 (dd, J = 5.7, 4.8 Hz, 2H), 3.74 – 3.68 (m, 2H), 3.68 – 3.63 (m, 2H), 3.59 (d, J = 5.1 Hz, 2H), 3.51 (t, J = 5.2 Hz, 2H), 2.86 (t, J = 5.2 Hz, 2H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 148.4, 136.1, 130.5, 130.2, 124.3, 114.0, 73.5, 70.8, 70.3, 68.5, 43.2, 41.7.
HRMS (ESI): m/z calc. for $C_{12}H_{19}N_4O_6$ 315.1299; found 315.1293, $[M+H]^+$.



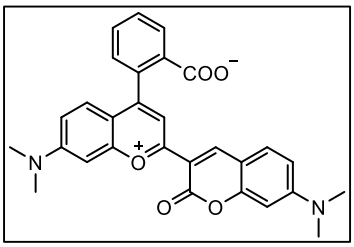
BC1-DN (9): Compound **8** (250 mg, 80.0 µmol) was added into a mixture of compound **4** (280 mg, 53.0 µmol), (benzotriazol-1-yloxy)tris(dimethylamino) phosphonium hexafluorophosphate (278 mg, 63.0 µmol) and N,N-diisopropylethylamine (20.0 mg, 155 µmol) in anhydrous dimethylformamide (1 mL). The mixture was stirred for 1 h. The mixture was then diluted with water (50 mL) and extracted with ethyl acetate (5 × 50 mL). The combined organic phases were dried over magnesium sulfate and evaporated under reduced pressure. The obtained blue solid was purified by RP-HPLC.
Yield: 5.30 mg, 6.46 µmol, 12%.
^{1}H NMR (400 MHz, CD_3OD) δ 8.93 (s, 1H), 8.87 (d, J = 2.8 Hz, 1H), 8.63 (d, J = 1.7 Hz, 1H), 8.23 (s, 1H), 8.21 – 8.14 (m, 2H), 7.61 (d, *J* = 9.0 Hz, 1H), 7.49 (d, J = 7.9 Hz, 1H), 7.25 – 7.07 (m, 4H), 6.92 (d, J = 9.0 Hz, 1H), 6.60 (s, 1H), 3.83 (t, J = 5.2 Hz, 2H), 3.73 (d, J = 4.8 Hz, 6H), 3.64 (t, J = 5.3 Hz, 2H), 3.60 (t, J = 5.2 Hz, 2H), 3.31 (s, 6H), 3.22 (d, J = 1.8 Hz, 6H).
^{13}C NMR (101 MHz, CD_3OD) δ 166.5, 166.5, 162.2, 162.1, 158.5, 158.4, 158.2, 157.3, 157.0, 148.3, 146.1, 139.0, 136.2, 135.6, 132.5, 130.9, 130.6, 130.0, 129.9, 129.7, 129.0, 123.3, 116.9, 116.0, 114.6, 111.9, 111.7, 110.0, 105.8, 96.6, 96.1, 70.2, 70.2, 69.1, 68.4, 42.8, 39.8, 39.7, 39.4.
HRMS (ESI): m/z calc. for $C_{42}H_{41}N_6O_{12}$ 821.2777; found 821.2783, $[M+H]^+$.

5.2.3 Synthesis of BC1-7

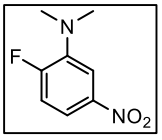


2-(4-(dimethylamino)-2-hydroxybenzoyl)benzoic acid (10): 3-Dimethylaminophenol (4.93 g, 36.0 mmol) was dissolved in toluene (10 mL) at 60 °C in a 100 mL flask. Powdered phthalic anhydride (5.33 g, 36.0 mmol) was added to this solution, and the mixture was heated under reflux for 24 h. The solvent was removed under reduced pressure. The residue was purified by column chromatography using cyclohexane:ethyl acetate mixtures (10:0 to 10:2).

Yield: 560 mg, 1.96 mmol, 6%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.10 (dd, J = 7.9, 1.2 Hz, 1H), 7.63 (td, J = 7.5, 1.4 Hz, 1H), 7.53 (td, *J* = 7.6, 1.4 Hz, 1H), 7.36 (dd, J = 7.5, 1.3 Hz, 1H), 6.89 (d, J = 9.1 Hz, 1H), 6.16 (d, J = 2.5 Hz, 1H), 6.07 (dd, J = 9.1, 2.5 Hz, 1H), 3.03 (s, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 198.6, 169.8, 165.2, 156.0, 141.1, 134.3, 132.8, 131.1, 129.2, 128.0, 127.6, 110.3, 104.0, 97.8, 40.0.
HRMS (ESI): m/z calc. for $C_{16}H_{16}NO_4$ 286.1074; found 286.1075, $[M+H]^+$.

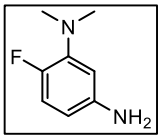


BC1 (11): Compounds **10** (456 mg, 1.60 mmol) and **3** (360 mg, 1.60 mmol) were dissolved in concentrated sulfuric acid (8 mL) and stirred for 6 h at 90 °C. The mixture was cooled down to room temperature, and ice (approx. 15 g) was carefully added. After adding 70% perchloric acid (0.70 mL), the mixture was filtered, and the precipitate was washed with water. The crude product was air-dried overnight and further purified by column chromatography using dichloromethane:methanol mixtures (50:1 to 20:1).
Yield: 477 mg, 1.00 mmol, 60%.
^{1}H NMR (500 MHz, C_2D_6OS) δ 9.10 (s, 2H), 8.13 (d, *J* = 7.8 Hz, 1H), 7.84 (td, *J* = 7.6, 1.3 Hz, 1H), 7.77 (t, *J* = 7.6 Hz, 1H), 7.73 (d, *J* = 9.1 Hz, 1H), 7.49 (d, *J* = 7.6 Hz, 1H), 7.20 (s, 3H), 6.98 (dd, *J* = 9.1, 2.4 Hz, 1H), 6.71 (d, *J* = 2.4 Hz, 1H), 3.26 (s, 6H), 3.19 (s, 6H).
^{13}C NMR (75 MHz, C_2D_6OS) δ 169.2, 158.9, 156.1, 154.3, 153.8, 152.2, 152.2, 151.5, 147.2, 141.9, 135.7, 130.8, 130.3, 128.8, 128.2, 127.3, 126.0, 125.0, 124.3, 110.6, 110.4, 108.4, 105.7, 98.6, 97.0, 40.3, 40.2.
HRMS (ESI): m/z calc. for $C_{29}H_{25}N_2O_5$ 481.1758; found 481.1747, $[M+H]^+$.

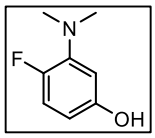


2-fluoro-5-nitro-N,N-dimethylaniline (12): 3 M Sulfuric acid (16.0 mL, 48.0 mmol) and 37% formaldehyde (19.0 mL, 234 mmol) were added into tetrahydrofuran (150 mL). A suspension of 2-fluoro-5-nitroaniline (9.98 g, 64.0 mmol) and sodium borohydride (14.0 g 370 mmol) in 150 mL tetrahydrofuran was added to the above solution in small portions. The reaction mixture was acidified with 3 M sulfuric acid (16.0 mL, 48.0 mmol) again when half of the suspension was added. After 15 min stirring, potassium hydroxide (1 M) was added to the reaction mixture until pH 10 was reached and the organic phase was separated. The aqueous phase was extracted with ethyl acetate (3 × 150 mL). The combined organic phases were washed with brine (2 × 200 mL), dried over magnesium sulfate, and evaporated under reduced pressure. Further purification was carried out by column chromatography using cyclohexane:ethyl acetate mixtures (10:0 to 10:0.5).
Yield: 3.10 g, 16.8 mmol, 26%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.70 (ddd, *J* = 6.9, 2.9, 0.9 Hz, 2H), 7.08 (ddd, *J* = 12.3, 9.8, 0.7 Hz, 1H), 2.94 (d, *J* = 1.3 Hz, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 157.9 (d, *J* = 256.3 Hz), 144.6, 141.2 (d, *J* = 10.0 Hz), 116.5 (d, *J* = 24.2 Hz), 115.7 (d, *J* = 9.4 Hz), 113.0 (d, *J* = 6.2 Hz), 42.3, 42.3.

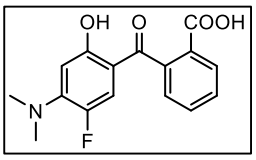
HRMS (ESI): m/z calc. for $C_8H_{10}FN_2O_2$ 185.0721; found 185.0723, $[M+H]^+$.



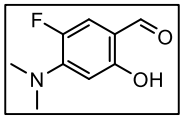
5-amino-2-fluoro-N,N-dimethylaniline (13): To a Schlenk flask filled with argon, 10% palladium on carbon (90 mg) and ethyl acetate (4 mL) were added. The argon was exchanged by hydrogen gas, and the mixture was stirred vigorously for 15 min at room temperature to activate the catalyst. Compound **12** (595 mg, 3.20 mmol) in ethyl acetate (3 mL) was added to the solution and stirred for 4 h at room temperature. After the reaction was completed, the hydrogen was replaced with argon. The mixture was filtrated through Celite, and the filter cake was washed thoroughly with ethyl acetate. The solvent was evaporated under reduced pressure to yield pure **13**.
Yield: 498 mg, 3.23 mmol, quantitative.
^{1}H NMR (300 MHz, $CDCl_3$) δ 6.77 (ddd, *J* = 15.8, 12.1, 8.5 Hz, 1H), 6.33 – 5.78 (m, 2H), 2.94 – 2.67 (m, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 149.0 (d, *J* = 235.4 Hz), 142.7 (d, *J* = 2.2 Hz), 141.1 (d, *J* = 9.9 Hz), 116.3 (d, *J* = 22.3 Hz), 107.0 (d, *J* = 7.5 Hz), 105.5 (d, *J* = 2.8 Hz), 42.8, 42.7.
HRMS (ESI): m/z calc. for $C_8H_{12}FN_2$ 155.0979; found 155.0981, $[M+H]^+$.



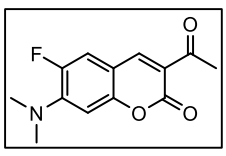
2-fluoro-5-hydroxy-N,N-dimethylaniline (14): Compound **13** (2.80 g, 18.0 mmol) was dissolved in 35% sulfuric acid (22 mL), and the temperature was decreased to 0 °C. Sodium nitrite (1.52 g, 22.0 mmol) dissolved in water (16 mL) was added, and the reaction mixture was stirred for 30 min. Following the addition of several crystals of urea, the mixture was poured into a solution of copper (II) nitrate trihydrate (60.0 g, 248 mmol) in water (600 mL). After stirring for several minutes, copper (I) oxide (2.15 g, 15.0 mmol) was added in small portions. The mixture was stirred vigorously for 1 h. Then, a saturated solution of sodium bicarbonate was added until pH 8 was reached. The suspension was filtered, and the residue was washed thoroughly with ethyl acetate. The filtered solution was extracted with ethyl acetate (3 × 200 mL). The combined organic phases were washed with brine, dried over magnesium sulfate, and evaporated under reduced pressure. The crude product was purified via column chromatography using cyclohexane:ethyl acetate mixtures (10:0 to 10:3).
Yield: 1.70 g, 11.0 mmol, 61%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 6.85 (dd, *J* = 12.7, 8.6 Hz, 1H), 6.38 (dd, *J* = 7.4, 3.0 Hz, 1H), 6.26 (dt, *J* = 8.6, 3.1 Hz, 1H), 2.82 (d, *J* = 0.9 Hz, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 151.8, 149.6 (d, *J* = 237.0 Hz), 141.4 (d, *J* = 10.2 Hz), 116.3 (d, *J* = 22.9 Hz), 106.4 (d, *J* = 7.8 Hz), 105.5 (d, *J* = 3.4 Hz), 42.7, 42.6.
HRMS (ESI): m/z calc. for $C_8H_{11}FNO$ 156.0819; found 156.0821, $[M+H]^+$.



2-(4-(dimethylamino)-5-fluoro-2-hydroxybenzoyl)benzoic acid (15): A mixture of powdered phthalic anhydride (296 mg, 2.00 mmol) and compound **14** (155 mg, 1.00 mmol) was heated at 150 °C under argon. After 20 min, zinc chloride (136 mg, 1.00 mmol) was added, and the reaction mixture was heated up to 180 °C for 6 h. After cooling to room temperature, water (10 mL) was added. The suspension was filtered, and the precipitate was washed twice with cold water. A dark violet solid was obtained. The solid was dissolved in 35% sodium hydroxide (5 mL) and stirred at 90 °C for 8 h under argon. After the reaction was complete, the mixture was neutralized with hydrochloric acid (conc.) to pH 6 in an ice bath. A dark pink precipitate was formed and collected via filtration. The crude product was purified by column chromatography with dichloromethane:methanol mixtures (10:0 to 10:1).
Yield: 69.0 mg, 227 µmol, 23%.
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.13 (dd, *J* = 7.9, 1.3 Hz, 1H), 7.67 (td, *J* = 7.5, 1.3 Hz, 1H), 7.58 (td, *J* = 7.7, 1.3 Hz, 1H), 7.36 (dd, *J* = 7.6, 1.3 Hz, 1H), 6.61 (d, *J* = 14.8 Hz, 1H), 6.26 (d, *J* = 7.8 Hz, 1H), 3.04 (d, *J* = 2.0 Hz, 6H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 198.8 (d, *J* = 2.4 Hz), 168.7, 160.8, 147.3 (d, *J* = 9.6 Hz), 145.5 (d, *J* = 236.8 Hz), 140.7, 132.9, 131.1, 129.5, 127.6, 127.6, 118.2 (d, *J* = 23.8 Hz), 109.8 (d, *J* = 6.5 Hz), 102.9 (d, *J* = 3.7 Hz), 42.1, 42.0.
HRMS (ESI): m/z calc. for $C_{16}H_{14}FNNaO_4$ 326.0799; found 326.0808, $[M+Na]^+$.



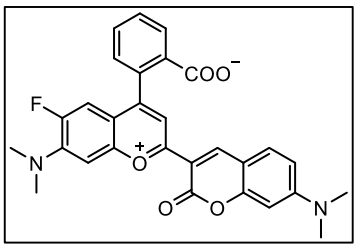
5-fluoro-4-(dimethylamino)-2-hydroxybenzaldehyde (16): Phosphorus(V) oxychloride (1.53 g, 10.0 mmol) was added dropwise into anhydrous dimethylformamide (15 mL) at 0 °C under argon. Compound **14** (605 mg, 3.90 mmol) in dimethylformamide (10 mL) was slowly added into this solution. The reaction mixture was brought to room temperature, stirred for 30 min, heated to 75 °C, and stirred for 2 h. After the reaction was complete, ice-cold water (30 mL) was added. The mixture was neutralized with sodium acetate and extracted with ethyl acetate (4 × 100 mL). The combined organic phases were washed with brine, dried over magnesium sulfate, and evaporated. The crude product was purified by column chromatography using cyclohexane:ethyl acetate mixtures (10:0 to 10:2)
Yield: 249 mg, 1.36 mmol, 36%.
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.04 (d, *J* = 13.8 Hz, 1H), 6.17 (d, *J* = 7.5 Hz, 1H), 3.07 (d, *J* = 2.2 Hz, 6H).
^{13}C NMR (75 MHz, $CDCl_3$) δ 192.0 (d, *J* = 2.3 Hz), 160.1, 147.5 (d, *J* = 9.4 Hz), 146.2 (d, *J* = 237.5 Hz), 118.6 (d, *J* = 23.0 Hz), 110.7 (d, *J* = 6.2 Hz), 102.2 (d, *J* = 3.7 Hz), 42.1, 42.0.
HRMS (ESI): m/z calc. for $C_9H_{10}FNNaO_2$ 206.0588; found 206.0592, $[M+Na]^+$.



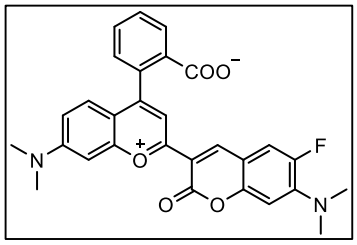
3-acetyl-7-(dimethylamino)-6-fluoro-2H-chromen-2-one (17): Compound **16** (238 mg, 1.30 mmol) and ethyl acetoacetate (182 mg, 1.40 mmol) were dissolved in absolute ethanol (3 mL). A few drops of piperidine were slowly added to the solution, and

the mixture was refluxed for 4 h. The reaction mixture was kept at 4 °C overnight, leading to a yellow precipitate formation, which was filtered off and washed with ice-cold ethanol. The crude product was purified with column chromatography using cyclohexane:ethyl acetate mixtures (10:0.5 to 10:2).
Yield: 241 mg, 0.97 mmol, 77%.
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.39 (s, 1H), 7.16 (d, *J* = 13.4 Hz, 1H), 6.58 (d, *J* = 7.6 Hz, 1H), 3.13 (d, *J* = 2.2 Hz, 7H), 2.69 (s, 3H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 195.5, 160.2, 154.2, 149.8 (d, *J* = 244.4 Hz), 147.0 (d, *J* = 3.1 Hz), 145.8 (d, *J* = 10.1 Hz), 118.9, 115.5 (d, *J* = 24.1 Hz), 108.6 (d, *J* = 9.4 Hz), 102.1 (d, *J* = 4.6 Hz), 42.3, 42.3, 30.6.
HRMS (ESI): m/z calc. for $C_{13}H_{12}FNNaO_3$ 272.0693; found 272.0702, $[M+Na]^+$.

The general method of the condensation reaction of benzoic acid and coumarin to BC2, BC3, and BC4: Compounds **15** (30.3 mg, 100 μmol) and **3** (23.0 mg, 100 μmol) for **BC2**, compounds **10** (28.5 mg, 100 μmol) and **17** (24.9 mg, 100 μmol) for **BC3**, compounds **15** (29.0 mg, 100 μmol) and **17** (24.9 mg, 100 μmol) for **BC3** were dissolved in concentrated sulfuric acid (2 mL) and stirred at 90 °C for 24 h under argon. The mixture was cooled down to room temperature, and crushed ice (approx. 10 g) was added. After adding 70% perchloric acid (2 drops), the mixture was filtered. The precipitate was washed with water, dried in the air overnight, and purified by RP-HPLC.

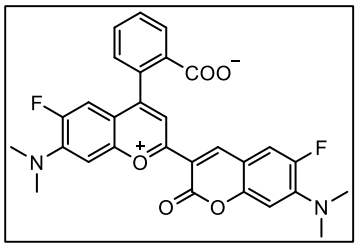
BC2:


Yield: 7.30 mg, 14.7 μmol, 15%.
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.41 (s, 1H), 8.03 (d, *J* = 7.7 Hz, 1H), 7.67 (td, *J* = 7.5, 1.2 Hz, 1H), 7.60 (td, *J* = 7.5, 1.1 Hz, 1H), 7.49 (d, *J* = 8.9 Hz, 1H), 7.27 (d, *J* = 5.9 Hz, 2H), 6.80 (d, *J* = 7.6 Hz, 2H), 6.69 (dd, *J* = 8.9, 2.5 Hz, 1H), 6.49 (d, *J* = 2.4 Hz, 1H), 6.39 (d, *J* = 13.3 Hz, 1H), 3.12 (s, 6H), 3.00 – 2.93 (m, 6H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 169.2, 158.9, 156.3, 154.0, 152.1, 150.2, 134.4, 130.2 (d, *J* = 3.1 Hz), 129.7, 126.9, 124.5 (d, *J* = 5.4 Hz), 113.5 (d, *J* = 23.9 Hz), 110.0, 108.8, 104.6, 97.1, 42.5 (d, *J* = 5.4 Hz), 40.3.
HRMS (ESI): m/z calc. for $C_{29}H_{24}FN_2O_5$ 499.1664; found 499.1662, $[M+H]^+$.

BC3:


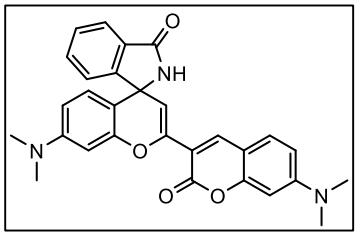
Yield: 7.20 mg, 14.5 μmol, 10%.
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.36 (s, 1H), 7.98 (dd, *J* = 7.5, 1.0 Hz, 1H), 7.64 (td, *J* = 7.5, 1.2 Hz, 1H), 7.57 (td, *J* = 7.5, 1.0 Hz, 1H), 7.27 – 7.22 (m, 2H), 6.73 (s, 1H), 6.63 (dd, *J* = 12.9, 8.2 Hz, 2H), 6.55 (d, *J* = 2.5 Hz, 1H), 6.43 (dd, *J* = 8.9, 2.6 Hz, 1H), 3.07 (d, *J* = 1.9 Hz, 6H), 3.01 (s, 6H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 169.6, 158.5, 152.6, 152.1, 151.4 (d, *J* = 55.1 Hz), 149.3, 144.6 (d, *J* = 10.2 Hz), 139.6, 134.4, 129.4, 128.4, 126.8, 125.3, 124.1, 114.3 (d, *J* = 24.4 Hz), 110.0, 109.6, 109.5, 102.8, 102.8, 98.3, 42.3, 42.2, 40.3.
HRMS (ESI): m/z calc. for $C_{29}H_{24}FN_2O_5$ 499.1664; found 499.1653, $[M+H]^+$.

BC4:
Yield: 5.70 mg, 11.0 µmol, 11%.


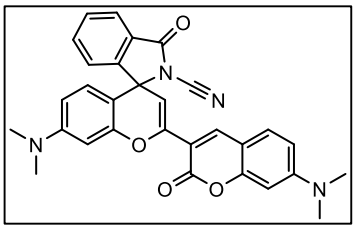
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.37 (s, 1H), 8.00 (dt, J = 7.4, 1.1 Hz, 1H), 7.62 (dtd, J = 22.6, 7.4, 1.2 Hz, 2H), 7.26 (d, *J* = 2.1 Hz, 1H), 7.25 – 7.21 (m, 1H), 6.78 (d, J = 7.5 Hz, 2H), 6.60 (d, J = 7.5 Hz, 1H), 6.36 (d, J = 13.3 Hz, 1H), 3.07 (d, J = 1.9 Hz, 7H), 2.95 (d, J = 1.3 Hz, 6H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 169.1, 158.4, 152.2, 151.9, 151.1, 150.2, 149.2, 144.8 (d, *J* = 10.6 Hz), 134.5, 129.8, 126.7, 124.3, 114.5 (d, *J* = 24.8 Hz), 113.5 (d, *J* = 23.7 Hz), 109.4, 104.6, 102.6 (d, *J* = 3.3 Hz), 42.5, 42.5, 42.3, 42.3.
HRMS (ESI): m/z calc. for $C_{29}H_{23}F_2N_2O_5$ 517.1570; found 517.1570, $[M+H]^+$.

The general method for the amidation of BC1 to BC5, BC6, and BC7. To a solution of **BC1** (23.3 mg, 50.0 µmol) in anhydrous dichloromethane (3 mL), phosphorus(V) oxychloride (61.0 mg, 400 µmol) was added dropwise. The mixture was refluxed for 4 h. After cooling, the reaction mixture was evaporated under reduced pressure to obtain the crude acid chloride. An ammonia solution (conc. 300 µL) for **BC5** or cyanamide (42.0 mg, 1.00 mmol) for **BC6** or methanesulfonamide (95.1 mg, 1.00 mmol) for **BC7** and triethylamine (40.0 mg, 396 µmol) in acetonitrile (2 mL) was added to the acid chloride. The reaction mixture was stirred at room temperature. After 2 h, water was added to the reaction mixture, and the aqueous phase was extracted with dichloromethane. The combined organic layers were dried over magnesium sulfate and purified with RP-HPLC.

BC5:
Yield: 4.30 mg, 8.98 µmol, 18%.


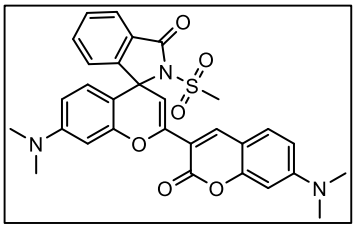
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.32 (s, 1H), 7.89 (d, *J* = 7.6 Hz, 1H), 7.54 – 7.43 (m, 3H), 7.21 (d, *J* = 7.2 Hz, 1H), 6.74 – 6.67 (m, 2H), 6.52 (dd, J = 12.7, 2.5 Hz, 2H), 6.44 (d, J = 8.6 Hz, 2H), 3.11 (s, 6H), 3.00 (s, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 169.3, 159.2, 155.8, 154.3, 154.3, 153.4, 151.7, 151.2, 145.0, 139.4, 132.8, 130.1, 129.5, 128.4, 128.2, 124.0, 123.3, 112.7, 109.6, 108.7, 107.7, 103.3, 98.8, 97.3, 58.3, 40.4, 40.2.
HRMS (ESI): m/z calc. for $C_{29}H_{26}N_3O_4$ 480.1918; found 480.1911, $[M+H]^+$.

BC6:



Yield: 3.80 mg, 7.54 µmol, 14%.
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (s, 1H), 7.89 (d, J = 7.7 Hz, 1H), 7.60 (td, J = 7.5, 1.2 Hz, 1H), 7.49 (t, *J* = 7.5 Hz, 1H), 7.39 (d, J = 8.8 Hz, 1H), 7.22 (d, J = 7.7 Hz, 1H), 6.61 (dd, J = 8.8, 2.4 Hz, 1H), 6.50 – 6.42 (m, 4H), 6.34 (dd, J = 8.9, 2.5 Hz, 1H), 3.03 (s, 6H), 2.92 (s, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 166.1, 159.2, 156.0, 153.7, 152.6, 151.9, 151.8, 147.9, 140.8, 135.4, 129.9, 129.6, 127.8, 126.4, 124.9, 124.6, 111.6, 110.1, 109.8, 108.6, 107.1, 103.7, 98.8, 98.7,
HRMS (ESI): m/z calc. for $C_{30}H_{24}N_4NaO_4$ 527.1690; found 527.1682, $[M+Na]^+$.

BC7:



Yield: 5.80 mg, 10.0 µmol, 20%.
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (s, 1H), 7.88 (d, J = 7.7 Hz, 1H), 7.55 (td, J = 7.5, 1.2 Hz, 1H), 7.45 (td, J = 7.5, 1.0 Hz, 1H), 7.38 (d, J = 8.8 Hz, 1H), 7.17 (d, J = 7.7 Hz, 1H), 6.60 (dd, J = 8.9, 2.4 Hz, 1H), 6.54 (d, J = 8.8 Hz, 1H), 6.44 (dd, J = 7.1, 2.4 Hz, 2H), 6.39 (s, 1H), 6.32 (dd, J = 8.8, 2.5 Hz, 1H), 3.14 (s, 3H), 3.03 (s, 6H), 2.91 (s, 6H).
^{13}C NMR (101 MHz, $CDCl_3$) δ 167.0, 159.6, 155.9, 153.5, 152.8, 152.6, 151.4, 146.8, 140.4, 134.9, 129.7, 129.2, 127.7, 127.5, 124.7, 124.2, 112.3, 109.7, 109.3, 108.8, 106.6, 100.4, 98.7, 97.2, 68.2, 42.5, 40.3, 40.2.
HRMS (ESI): m/z calc. for $C_{30}H_{27}N_3NaO_6S$ 580.1513; found 580.1510, $[M+Na]^+$.

5.2.4 Preparative RP-HPLC

The preparative RP-HPLCs were performed on an Agilent 1100 series HPLC equipped with a 5 µm C18 Phenomenex Luna reversed-phase column at a flow rate of 5 mL/min. The following elution solvents were used: solvent A (0.1% trifluoroacetic acid in water (v/v)), solvent B (0.1% trifluoroacetic acid in acetonitrile (v/v)), solvent C (100 mM triethylammonium acetate in water, pH 7.0) and solvent D (100 mM triethylammonium acetate in acetonitrile/water (8/2)). The HPLC-purified fractions were dried on an Alpha 1-4 LD lyophilizer.

Compound	Elution method	Retention time
BC1-COOH	0 min, 60% A, 40% B; 30 min, 50% A, 50 %B	26 min
BC1-NH_2	0 min, 70% C, 30% D; 40 min, 55% C, 45 %D	24 min
BC1-biotin	0 min, 65% A, 35% B; 40 min, 65% A, 35 %B	34 min
BC1-DN	0 min, 50% C, 50% D; 40 min, 50% C, 50 %D	35 min
BC1	0 min, 55% A, 45% B; 50 min, 55% A, 45 %B	39 min
BC2	0 min, 63% A, 37% B; 40 min, 63% A, 37 %B	22 min
BC3	0 min, 50% A, 50% B; 40 min, 50% A, 50 %B	27 min
BC4	0 min, 50% A, 50% B; 40 min, 50% A, 50 %B	34 min
BC5	0 min, 50% A, 50% B; 40 min, 40% A, 60 %B	38 min
BC6	0 min, 40% A, 60% B; 45 min, 10% A, 90 %B	40 min
BC7	0 min, 50% A, 50% B; 50 min, 50% A, 50 %B	44 min

Table 18. Preparative RP-HPLC elution conditions for all the fluorophores. The elution gradient and the retention time for each fluorophore are shown.

5.3 Aptamer selection

For molecular biology experiments, DNA and RNA concentrations were determined using a NanoDrop ND-1000 spectrophotometer. Synthetic DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). Sequencing was carried out at Microsynth Seqlab.

5.3.1 *In vitro* aptamer selection for BC zwitterion

The aptamer selection protocol described in this chapter was mostly covered in the electronic supplementary of an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.

5.3.1.1 DNA library preparation

The DNA library contains a T7 promoter sequence, two fixed prime-binding sites, two 26-nt random regions (indicated with N) interspersed by a 12-nt constant region (Table 19). A reverse complementary ssDNA library (1 µmol) was ordered from IDT. The phosphoramidites used to synthesize the random regions were hand-mixed at an A:C:G:T ratio of 3:3:2:2. The dsDNA library was prepared by PCR amplification (6 cycles) using the ssDNA as templates (3×10^{15} molecules). PCR conditions: primer A (10 µM), primer B (10 µM), dNTPs (1 mM), $MgCl_2$ (6 mM), *Taq* DNA polymerase, and 1 × PCR buffer. The PCR product was PCI purified and used for *in vitro* transcription of the RNA library.

DNA library (122-nt)	5′ – TC TAA TAC GAC TCA CTA TA GGA GCT CAG CCT TCA CTG C NNN NNN NNN NNN NNN NNN NNN NNN NN CTG CTT CGG CAG NNN NNN NNN NNN NNN NNN NNN NNN NN GGC ACC ACG GTC GGA TCC AC – 3′ (Underline: T7 promoter; Green: primer-binding site; Blue: inserted constant region)
ssDNA template (99-nt)	5′ – GAT CCG ACC GTG GTG CC NNN NNN NNN NNN NNN NNN NNN NNN NN CTG CCG AAG CAG NNN NNN NNN NNN NNN NNN NNN NNN NN GCA GTG AAG GCT GAG TCC – 3′
Primer A	5′ – T CTA ATA CGA CTC ACT ATA GGA GCT CAG CCT TCA CTG C-3′
Primer B	5′ - GTG GAT CCG ACC GTG GTG CC – 3′

Table 19. The sequences of the DNA library, ssDNA template, primer A and B.

5.3.1.2 Resin preparation

Sepharose resin functionalized with BC1-NH_2 for selection rounds 1–5.

4 mL of NHS-activated sepharose 4 Fast Flow resin was transferred into a 15 mL falcon tube and washed with 1 mM hydrochloric acid (2 × 8 mL, ice-cold). The resin was resuspended in 1 mL of HEPES buffer (100 mM, pH 7.4). **BC1-NH_2** (11 mg, 17 µmol) was dissolved in 100 µL of dimethyl sulfoxide (DMSO) and further diluted with 1 mL of HEPES buffer (100 mM). The dye solution was added to the resin on ice and mixed thoroughly. The suspension was kept at 25 °C overnight with vigorous shaking. The next day, the resin was further incubated with ethanolamine (0.2 M, 12 µL) at 25 °C for 2 h to quench any unreacted NHS groups. The resin was washed with HEPES buffer (50 mM, 6 × 8 mL) until the supernatant became colorless. The resin was washed once more with NaAc buffer (0.1 M, pH 5.5) and stored in NaAc buffer (0.1 M, pH 5.5) containing 20% ethanol at 4 °C. The coupling efficiency of **BC1-NH_2** was determined to be 4 mM by measuring the absorption, thereby the concentration of the dye before reaction (total dye) and after reaction in wash solutions (unreacted dye). The mock resin for the negative selection was prepared analogously without adding **BC1-NH_2**.

Streptavidin-agarose resin functionalized with BC1-biotin for selection rounds 6–11.

The resin for selection was switched to the streptavidin-functionalized agarose resin to regulate the **BC1** concentration on beads down to a low µM level. **BC1-biotin** (7.5 mM in DMSO) was diluted with 200 µL of 1 × ASB and loaded to the streptavidin resin in portions with vigorous shaking. After incubating for 5 min, the beads were washed 6 times with 1 × ASB (1 mL). For the negative selection, the streptavidin beads were used directly without functionalization.

5.3.1.3 Selection protocol

The RNA library for the first selection round was prepared by large-scale *in vitro* transcription reaction (10 mL) using the dsDNA library (1.5 µM) as templates and purified by PCI extraction. **(1) Mock selection.** The folded RNA (round 1: 100 µM, round 2: 50 µM, round 3–11: 10 µM) was first incubated with mock resin (round 1: 500 µL, round 2-11: 250 µL) at 25 °C for 1 h to remove resin-binding RNA. **(2) RNA binding.** The remaining RNA was incubated with the **BC1** conjugated resin at 25 °C for 30 min. The resin was then washed with 1 × ASB (round 1: 6 column volumes (cv), rounds 2–5: 12 cv, rounds 6–11: 30 cv). **(3) RNA elution.** The bound RNA was eluted with 98% formamide containing 10 mM EDTA at 65 °C for 5 min (rounds 1–5) or 50 mM DTT at 25 °C for 10 min (rounds 6–11). The eluted RNA was co-precipitated with glycogen (1 µg/µL) using ethanol. The RNA pellet was rinsed twice with 75% ethanol and dissolved in ultrapure water. **(4) Reverse transcription and PCR amplification.** The RNA was reverse transcribed (RT) by SuperScript IV reverse transcriptase (reaction volume: 200 µL) according to the manufacturer's instructions. The RT reaction was immediately used for PCR amplification (800 µL, 10 cycles). **(5) RNA *in vitro* transcription.** RNA was transcribed from the DNA template (reaction volume

for rounds 2–11: 800 µL), purified by 10% PAGE, and used as input RNA for the next selection round. The selection was carried out for 11 iterative rounds using the described protocol and monitored by calculating the fraction of eluted RNA based on UV-spectrometric measurement (Nanodrop).

$$Eluted\ RNA\ fraction\ (\%) = \frac{Number\ of\ eluted\ RNA\ molecules}{Number\ of\ input\ RNA\ molecules}$$

5.3.2 Fluorescence-activated bacteria sorting for SiRA evolution

5.3.2.1 DNA library preparation

The enriched DNA pools (round 7 and 14) from SELEX against SiR were provided by Dr. Regina Wirth. Both DNA libraries (~ 1 µg of DNA template for each library) were PCR amplified (10 cycles) using the sequence-specific primers containing the SalI and AatII recognition sequences, respectively (Table 20). PCR conditions: forward primer (5 µM), reverse primer (5 µM), dNTPs (1 mM), $MgCl_2$ (6 mM), *Taq* DNA polymerase, and 1 × PCR buffer.
The 100-nt ssDNA doped SiRA library (1 µmol) was ordered from IDT. It contains two fixed primer-binding sites (25-nt for each) and a partially randomized (doped) region (50-nt). Each primer-binding site contains the respective recognition sequences (SalI and AatII, respectively). The phosphoramidites used for the partially random region synthesis were hand-mixed in a controlled manner, where a doping rate of 9% based on the sequence of SiRA was applied (i.e., 91% wild-type residues and 3% of each non-wild type residue). The dsDNA library was obtained by PCR amplification (6 cycles) using the ssDNA library as templates (2 × 10^{13} molecules). PCR conditions: forward primer (5 µM), reverse primer (5 µM), dNTPs (1 mM), $MgCl_2$ (6 mM), *Taq* DNA polymerase, and 1 × PCR buffer.

Enriched DNA pools	DNA template	5′ - GGA GCT CAG CCT TCA CTG C N_{26} CTG CTT CGG CAG N_{26} GGC ACC ACG GTC GGA TCC AC – 3′
	Forward primer	5′ – AAA AAA GTC GAC GGA GCT CAG CCT TCA CTG C - 3′
	Reverse primer	5′ – AAA AAA GAC GTC GTG GAT CCG ACC GTG GTG CC – 3′
Doped SiRA library	DNA template	5′ – G CTC AGC CTT CAC TGC GTC GAC GGT (N3)(N3)(N2) (N2)(N1)(N2) (N2)(N3)(N3) (N3)(N4)(N4) (N4)(N3)(N1) (N1)(N1)(N1) (N2)(N2)(N4) (N3)(N3)(N2) (N4)(N3)(N2) (N4)(N4)(N2) (N3)(N3)(N2) (N1)(N3)(N4) (N4)(N3)(N4) (N1)(N4)(N2) (N2)(N4)(N4) (N4)(N3)(N3) (N2)(N2) GAC GTC GAT GGC ACC ACG GTC GGA T – 3′
	Forward primer	5′ – T CTA ATA CGA CTC ACT ATA GGA GCT CAG CCT TCA CTG C -3′
	Reverse primer	5′ - GTG GAT CCG ACC GTG GTG CC – 3′

Table 20. DNA libraries for SiRA evolution. The sequences of enriched DNA pools (round 7 and 14) from SELEX against SiR, a 9% doped SiRA library, and their respective primers are shown. SalI and AatII sequences are indicated in blue and green, respectively. N represents a nucleotide with the probabilities in the order of ACGT (N: 30302020; N1: 91030303; N2: 03910303; N3: 03039103; N4: 03030391).

5.3.2.2 tRNA vector preparation

The SalI and AatII restriction sites were inserted into the anticodon region of a $tRNA_{Lys}$ scaffold within the *pET-tRNA*$_{Lys}$ plasmid via overhang PCR[56]. The PCR was carried out using Q5 high-fidelity DNA polymerase with sequence-specific primers according to the manufacturer's instructions (Table 21). The linear product (0.25 pmol) was phosphorylated (T4 polynucleotide kinase) and purified. 50 ng of the DNA product was self-ligated (T4 DNA ligase) in 1 × ligation buffer (50 µL). Afterward, 4 µL of the ligation mixture was used for transformation, and the transformants were placed on an LB-kanamycin agar plate. Single colonies were picked and grown in LB medium. Plasmids were purified from the bacterial cultures (GeneJET plasmid miniprep kit) and sent for sequencing. The sequencing results confirmed the successful construction of the *pET-tRNA-2* plasmid.

Forward primer	5′ – T GACGTCGATGCGGGTCCAGGGTTCAAGTC - 3′
Reverse primer	5′ – AAACCGTCGACGCTGCTCTACCGACTGAGC - 3′
pET-tRNA-2	…GCC CGG ATA GCT CAG TCG GTA GAG CAG C GTC GAC GGT TTT GAC GTC GAT GCG GGT CCA GGG TTC AAG TCC CTG TTC GGG CGC CA… (Orange: tRNA scaffold; Blue: Sal I; Green: AatII)

Table 21. The sequences of the forward primer, reverse primer, and tRNA-2 in the pET-tRNA-2 plasmid.

5.3.2.3 Plasmid library preparation

The DNA pools (SELEX rounds 7 and 14, and doped SiRA, chapter 5.3.2.1) and the *pET-tRNA-2* expression vector were treated with SalI and AatII in 1 × digestion buffer at 37 °C. The purified vector and inserts were mixed at a molar ratio of 1 : 10 and incubated with T4 DNA ligase at 22 °C overnight. The ligation mixtures were purified to yield L1, L2, and L3, respectively.

5.3.2.4 Fluorescence-activated bacteria sorting

The FACS-based aptamer selection was carried out using plasmid libraries L1, L2, and L3 as independent input. **(1) Bacteria library preparation.** The plasmid libraries (~ 200 ng for each plasmid pool) and the control plasmid (i.e., *pET-tRNA*) were transformed into competent BL21 cells (40 µL) via electroporation. **(2) Aptamer expression in bacteria.** For each group (L1, L2, L3, and control), the rescued bacteria cells after electroporation were grown in LB medium (15 mL, 30 µg/mL kanamycin) in a baffled flask until OD_{600} reached 0.4. After the addition of IPTG (final concentration: 1 mM), the bacteria were cultured for another 3 h (170 rpm, 37 °C). **(3) Sample preparation.** 700 µL of the bacteria culture from each

group was removed and placed into a new microcentrifuge tube. 100 µL of 8% paraformaldehyde (PFA) was added to the suspension. The suspension was incubated at room temperature for 10 min to allow fixation. The bacteria mixture was then pelleted by centrifugation (6000 rpm, 2 min), washed twice with M9 medium, and resuspended in M9 medium (1 mL). A few microliters of the fixed bacteria were diluted with 1 mL of M9 medium (filtered) to reach a final OD_{600} of 0.02. The desired amount of SiR-PEG_3-NH_2 (1 mM stock in DMSO) was added (dye concentration: 1 µM for rounds 1-3 and 200 nM for round 3L), and the mixture was incubated at 37 °C for 30 min. **(4) Bacteria sorting.** The fixed cells from each group were subjected to the cell sorter. A 633 nm laser was used to activate the fluorophore. Based on the fluorescence intensity, a sorting gate placed just above the control cells was applied to each experimental group for selection rounds 1 and 2. For round 3 and round 3-L, only the top 1% and 0.5% of the brightest cells were collected, respectively. A total of 10^4 events were collected for each sample. **(5) PCR amplification and cloning.** The isolated cells were washed once with ultrapure water and used as templates for PCR amplification. The amplified aptamer sequences were then cloned into the tRNA expression vector as described above (chapter 5.3.2.3). The generated plasmid constructs were used as input for the next round of sorting.

5.3.3 Sanger sequencing

The RNAs eluted from the final selection rounds (selections for BC and SiR, respectively) were reverse transcribed and PCR amplified. The purified DNA sequences were phosphorylated and ligated into a pDisplay-AP-CFP-TM plasmid that had already been digested with EcoRV and dephosphorylated[k]. The vectors were transformed into competent DH5α cells. Single colonies were picked and sent for Sanger sequencing.

5.4 *In vitro* spectroscopic characterization

The *in vitro* characterization of the fluorophores or the aptamer-fluorophore complexes described in this chapter was mostly covered by the electronic supplementary of an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.
All the fluorophores were stocked in DMSO at − 20 °C and diluted to the desired concentrations before measurement (final DMSO < 2%). RNAs were *in vitro* transcribed, PAGE purified, and folded *in vitro* before subjecting to the measurement unless noted otherwise.

[k] This EcoRV digested and dephosphorylated linear pDisplay-AP-CFP-TM plasmid was produced and kindly provided by Dr. Murat Sunbul.

5.4.1 Absorbance spectrum measurement

Working solutions were prepared by diluting the fluorophores to the desired concentrations in the required solvents (including organic solvents, buffers, and RNA-containing solutions). The concentration of the fluorophore and the composition of the measuring solvent were denoted in the respective figure legends for each experiment. 70 µL of the solution was placed in a quartz cuvette (light path: 10 mm), and the absorption spectrum was recorded on a Cary 50 UV-Visible spectrophotometer at room temperature. The scanning wavelength range was 300 - 800 nm with a medium scanning speed.

5.4.2 Excitation and emission spectrum measurement

Working solutions were prepared as described in chapter 5.4.1. 50 µL of the solution was placed in a quartz cuvette (light path: 3 × 3 mm), and the fluorescence measurement was performed on an FP-6500 fluorescence spectrometer with a slit width of ± 5 nm at room temperature unless noted otherwise. The scanning speed was 200 nm/min, and the data pitch was 1 nm. The excitation wavelength for each experiment was indicated in the respective figure legends.

5.4.3 Dissociation constant measurement

The RNA aptamer was folded as a stock solution in 1 × ASB and diluted with 1 × ASB to the desired concentrations. The fluorophore was then incubated with varying RNA aptamer concentrations, and the fluorescence of the mixtures was measured with a slit width of ± 5 nm. The maximal fluorescence intensity as a function of RNA concentration was plotted and fitted using the least-square fitting according to the given equation unless noted otherwise. The K_D value was determined accordingly in Origin. The ligand concentration and the excitation wavelength for each experiment were indicated in the respective figure legends.

$$F = F_0 + \frac{(F_\infty - F_0) \times \left\{(K_D + P_0 + [Apt]) - \sqrt{([Apt] - P_0)^2 + K_D \times (K_D + 2[Apt] + 2P_0)}\right\}}{2P_0}$$

F: fluorescence intensity at any given RNA concentration; F_0: fluorescence of a free probe with an initial concentration of P_0; F_∞: maximum fluorescence of the complex; $[Apt]$: concentration of added RNA; K_D: dissociation constant.

5.4.4 Dielectric constant titration of BC1-7

Solutions of **BC1-7** (5 µM) in water-dioxane mixtures containing 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% of dioxane (by volume) were prepared and the absorbance spectra were recorded. The absorbance of the zwitterionic form of **BC1-7** at each dielectric constant condition was normalized to that of each probe in water containing 0.3% sodium dodecyl sulfate (SDS) in which the

fluorophores were predominantly converted to the corresponding zwitterions. One exception was **BC5**, which predominately exists as spirolactam in water and could not be converted to its corresponding zwitterion by SDS. Thus, the absorbance of the **BC5** at each dielectric constant was normalized to that of its parent structure **BC1** in water containing 0.3% SDS. The normalized absorbance of **BC1-7** was plotted against the dielectric constant of the water-dioxane mixture[150].

5.4.5 Temperature and cation dependence measurement of BeCA-BC6

The emission spectra were recorded based on the standard method described in chapter 5.4.2.

5.4.5.1 Temperature dependence

BeCA (5 µM) was incubated with **BC6** or **BC1-DN** (5 µM) in 1 × ASB. The emission spectra were recorded as the temperature increased from 20 °C to 70 °C (2 °C interval in 2 min).

5.4.5.2 Magnesium and potassium dependence

BeCA (5 µM) was folded in either 20 mM HEPES (pH 7.4) buffter containing 125 mM KCl and varying concentations of $MgCl_2$ (0, 10, 25, 50, 100, 250, 500, 1000, 2500, 5000, 10000 µM) or 20 mM HEPES buffter (pH 7.4) containing 5 mM $MgCl_2$ and varying concentations KCl (0, 1, 2.5, 5, 10, 25, 50, 100, 200, 300 mM). Emission spectra were recorded upon the addition of **BC6** (5 µM).

5.4.6 Extinction coefficient and quantum yield determination of BC6 and BeCA-BC6

The absorbance and emission spectra of the free **BC6** (0.2, 0.4, 0.6, 0.8, 1 µM) or the **BeCA-BC6** complexes prepared by mixing varying concentrations of **BC6** (0.2, 0.4, 0.6, 0.8, 1 µM) with an excess amount of **BeCA** (25 µM) in 1 × ASB containing 0.05% Triton X-100 were recorded**.** Coumarin 153 and oxazine 1 in ethanol were used as the standard references for the quantum yield determination of the spirocyclic and zwitterionic **BC6**, respectively[151]. The excitation wavelengths for the cyan and NIR windows were 420 nm (± 3 nm slit width) and 615 nm (± 3 nm slit width), respectively.

Molar extinction coefficient

The maximal absorbance of each sample was plotted against the fluorophore's concentration, and the slope was determined using a linear fit. The molar extinction coefficient was calculated using the following equation:

$$A = \varepsilon \times c \times L$$

A: absorbance; ε: molar extinction coefficient; c: molar concentration of the probe; L: path length.

Quantum yield

The integrated area under the fluorescence emission spectrum was calculated and plotted against the sample's absorbance at the excitation wavelength. The slope was determined using a linear fit. The quantum yield was calculated using the following equation:

$$\emptyset = \emptyset_R \times \frac{S}{S_R} \times \frac{\eta^2}{\eta_R^2}$$

Φ: quantum yield; S: the slope of the fitted line; η: refractive index of the solvent; R: reference fluorophore with known quantum yield.

5.5 Live-cell RNA imaging

The live-cell RNA imaging described in this chapter was mostly covered in the electronic supplementary of an article "a color-shifting near-infrared fluorescent aptamer-fluorophore module for live-cell RNA imaging" by Zhang *et al*. (in preparation). The text was originally written by myself.

5.5.1 Construction of plasmids

All the digestions (restriction enzymes), phosphorylations (T4 polynucleotide kinase), dephosphorylations (FastAP thermosensitive alkaline phosphate), purifications (QIAquick PCR purification kit), and the ligations (T4 DNA ligase) were performed according to the manufacturers' instructions. Plasmids were purified from bacterial cultures using GeneJET plasmid miniprep kit and sequenced at Microsynth Seqlab.

5.5.1.1 Cloning of *pET-gfp-BeCA$_8$*

A 182-nt ssDNA ultramer was ordered from IDT. It contains 4 synonymous repeats of *BeCA* flanked by the SalI and XhoI recognition sites at 5'- and 3'-ends, respectively (Table 22). The ultramer (10 pmol) was PCR amplified (20 cycles) using the standard protocol. 1 µg of the purified DNA product was double-digested by SalI and XhoI at 37 °C overnight. Analogously, 1GFP plasmid (1 µg) was digested by SalI and dephosphorylated. Both DNA fragments were purified and ligated to the backbone on a scale of 10 µL with a molar ratio of 1 : 10 (vector : insert). The obtained plasmid containing 4 repeats of *BeCA* was further digested by XhoI, dephosphorylated, and ligated again with the double-digested PCR product using the same protocol. Afterward, 4 µL of the ligation mixture was used for transformation, and the transformants were placed on an LB-kanamycin agar plate. Single colonies were picked and grown in LB medium. The plasmids

were purified from the bacterial cultures and sent for sequencing. The sequencing result confirmed the successful insertion of 8 repeats of *BeCA* to the vector backbone, yielding the desired *pET-gfp-BeCA*$_8$ plasmid.

DNA ultramer	AAAGTCGACGGAGGTACGTGGACTGCTTATGCAGTAGGAAGGCC TCCAAAAGGTCCTACGTGGACTGCTTATGCAGTAGGAAGGGGAC CAAAAGCTGCTACGTGGACTGCTTATGCAGTAGGAAGGGCAGCA AAACGAGCTACGTGGACTGCTTATGCAGTAGGAAGGGCTCGCTC GAGAAA (Blue: SalI; Green: XhoI, Red: *BeCA*)
pET-gfp-BeCA$_8$	…GGAGGTACGTGGACTGCTTATGCAGTAGGAAGGCCTCCAAAA GGTCCTACGTGGACTGCTTATGCAGTAGGAAGGGGACCAAAAGC TGCTACGTGGACTGCTTATGCAGTAGGAAGGGCAGCAAAACGAG CTACGTGGACTGCTTATGCAGTAGGAAGGGCTCGCTCGACAAGC TTGCGGCCGCACTCGACGGAGGTACGTGGACTGCTTATGCAGTA GGAAGGCCTCCAAAAGGTCCTACGTGGACTGCTTATGCAGTAGG AAGGGGACCAAAAGCTGCTACGTGGACTGCTTATGCAGTAGGAA GGGCAGCAAAACGAGCTACGTGGACTGCTTATGCAGTAGGAAG GGCTCG… (Red: *BeCA*)
Forward primer	AAAGTCGACGGAGGTAC
Reverse primer	TTTCTCGAGCGAGCC

Table 22. The sequences of DNA ultramer, BeCA$_8$ *in the pET-gfp-BeCA*$_8$ *plasmid, the forward primer, and the reverse primer.*

5.5.1.2 Cloning of *pET-tRNA-BeCA*

The sequence encoding *BeCA* was introduced to the *tRNA* scaffold within the *pET-tRNA*$_{Lys}$ plasmid by PCR amplification using Q5 high-fidelity DNA polymerase (Table 23)[58]. The PCR product was phosphorylated, purified, and self-ligated to yield the *pET-tRNA-BeCA* plasmid using the method described above (chapter 5.3.2.2). Sanger sequencing confirmed the successful insertion of *BeCA* to the anticodon region of the tRNA scaffold.

pET-tRNA-BeCA	…TAATACGACTCACTATAGCGGATTTAGCTCAGTTGGGAGAGCGC CAGGGAGGTACGTGGACTGCTTATGCAGTAGGAAGGCCTCCCTGG AGGTCCTGTGTTCATCCACAGAATTCGCACCA… (Blue: T7 promoter; Green: tRNA scaffold; Red: *BeCA*)
Forward primer	ATGCAGTAGGAAGGCCTCCCTGGAGGTCCTGTGTTCG
Reverse primer	AAGCAGTCCACGTACCTCCCTGGCGCTCTCCCAACTG

Table 23. The sequences of tRNA-BeCA in the pET-tRNA-BeCA plasmid, the forward primer, and the reverse primer.

5.5.2 Live-bacteria confocal imaging using BeCA-BC6

BL21 Star (DE3) competent *E. coli* cells were transformed with the plasmid of interest (*1GFP*, *pET-gfp-BeCA*$_8$, *pET-tRNA* or *pET-tRNA-BeCA*) and grown overnight on an LB agar plate (30 µg/mL kanamycin) at 37 °C. A single colony was picked and grown in 5 mL of LB medium at 37 °C overnight with shaking

(170 rpm). The next day, a fresh culture was started with an OD_{600} of 0.02 in 10 mL of LB medium. When OD_{600} reached 0.4, the desired amount of IPTG was added to the medium to initiate transcription (170 rpm, 37 °C, 3 h). Then, 300 µL of the bacterial culture was removed, placed into a new microcentrifuge tube, spun down (6000 rpm, 2 min), washed twice with M9 medium, and resuspended in M9 medium (1 mL). The suspension was transferred into a poly-D-lysine (PDL) coated 8-well glass chamber. After cell attachment (room temperature, 15 min), the cells were washed twice with M9 medium and incubated with the desired amount of **BC6** (in 300 µL of M9 buffer containing 0.1% DMSO) for 15 min at 37 °C. Images were taken on a point scanning confocal microscope equipped with a hybrid-scanner (galvano/resonant) and a Nikon N Apo 60 × NA 1.4 λs OI objective (WD 0.14 mm, FOV 0.21 × 0.21 mm). Lasers of 405 nm (emission filter: 450 ± 50 nm), 488 nm (emission filter: 515 ± 30 nm), and 640 nm (emission filter: 700 ± 75 nm) were used for unbound **BC6**, GFP, and aptamer bound **BC6**, respectively. Images were analyzed in ImageJ/Fiji. Background correction was conducted by subtracting the mean fluorescence intensity of a surface area where no cells were present from the whole image.

5.5.3 Total RNA isolation and gel analysis of *tRNA-BeCA*

The *pET-tRNA-BeCA* plasmid was transformed into competent BL21 cells. The bacteria were grown and treated with different IPTG amounts (0, 10, 20, 50, 100, 1000 µM) as described above (chapter 5.5.2). After 3 h of induction, 3 mL of the bacterial culture from each IPTG-treated group was removed, pelleted, washed, and resuspended in M9 medium (600 µL). The bacterial suspensions were normalized to the same number according to the OD_{600} values, and 400 µL of the cell culture was used for total RNA isolation. Total RNA was isolated using RNAzol RT according to the manufacturer's instructions. Precipitated RNA was dissolved in water (100 µL), and 10 µL of the solution was loaded on a 10% PAGE gel for analysis.

5.6 Bacteria sorting

Selective media or agar plates (30 µg/mL kanamycin or 100 µg/mL ampicillin) were used to grow transformants carrying the respective antibiotic resistance genes. The bacterial cultures were grown in an incubator at 37 °C with shaking (170 rpm). All the analysis and sorting of bacteria in this thesis were performed in a FACS Aria III. For the RhoBAST-TMR-DN system, extra $MgCl_2$ was supplemented in the M9 medium (final magnesium concentration: 5 mM) to achieve the best fluorescence performance.

5.6.1 Bacteria sample preparation for method development

The plasmid of interest (*1GFP*, *pET-tRNA*, *pET-RhoBAST*, *pET-SiRA*, or *pTAC-GFP*) was transformed into competent *E. coli* cells (BL21 or DH5α). A single colony was picked from the agar plate, seeded into LB medium (5 mL), and

grown overnight. The next day, a fresh culture in 10 mL of LB was started with an OD_{600} of 0.02. When OD_{600} reached 0.4, the desired amount of IPTG was added to the medium. The bacteria were allowed to grow for another 3 h.

Live-cell sorting

500 µL of the bacterial culture was removed, spun down (6000 rpm, 2 min), washed with M9 medium, and resuspended in M9 medium (1 mL). The desired quantity of bacteria was diluted in 1 mL of M9 medium (filtered) with a final OD_{600} of 0.02, supplemented with the staining fluorophore (if needed, 30 min at 37 °C), and immediately subjected to a cell sorter.

Fixed-cell sorting

The bacteria (700 µL) were removed, diluted with 8% PFA (100 µL), and incubated at room temperature for 10 min. The bacteria mixture was then pelleted by concentration (6000 rpm, 2 min), washed twice with M9 medium, resuspended in M9 medium (1 mL), and stored at 4 °C. The desired quantity of fixed bacteria was diluted in 1 mL of M9 medium (filtered) with a final OD_{600} of 0.02. The staining fluorophore was added (if needed, 30 min at 37 °C) and sent to a cell sorter.

5.6.2 Instrument setup and operation

FSC and SSC in height (FSC-H and SSC-H) were set as the x and y axes, respectively, in the logarithmic scale. The threshold was adjusted to a value that an electronic background of 200 events/sec was observed by applying phosphate-buffered saline (PBS) buffer only. An FSC threshold of 2500 and an SSC threshold of 200 were proven sufficient and used for all the sorting experiments. By adjusting the sample concentration, a flow rate of 1 and not more than 3000 events/sec were applied for the sorting. Bacteria mixture with an OD_{600} of 0.02 was proven to be a proper concentration. Lasers of 488 nm, 561 nm, and 633 nm were used for the excitations of GFP (BP filter: 530 ± 30 nm), TMR-DN (BP filter: 582 ± 15 nm), and SiR (BP filter: 660 ± 20 nm), respectively. Based on the fluorescence signal, an appropriate sorting gate was applied for each sample, and 10000 events were collected for further use.

5.6.3 Cytotoxicity assay

The cell viability after IPTG treatment was determined by dilution plating. First, the bacteria transformed without or with a plasmid (*1GFP*, *pET-RhoBAST*, or *pTAC-GFP*) were grown to an OD_{600} of 0.4 and treated with varying concentrations of IPTG for 3 h using the same method described in chapter 5.6.1. Then, 200 µL of the bacteria culture from each IPTG-treated group was removed, and the cell density was normalized to an OD_{600} value of 1 with LB medium. This stock solution was termed Solution-0 (Figure 80). 20 µL of Solution-0 was mixed

with 180 µL of LB medium to generate Solution-1 with a dilution factor of 10. Analogously, this 10-fold dilution procedure was repeated four more times, yielding Solution-2, Solution-3, Solution-4, and Solution-5. 5 µL of the bacteria suspension from each dilution (Solutions 1-5) was placed on the agar plate with three replicates. The next day, colonies on the plate were manually counted, back-calculated with the dilution factors. The determined cell viability (CFU/mL) of each IPTG-treated group was normalized to that of the bacteria without IPTG treatment.

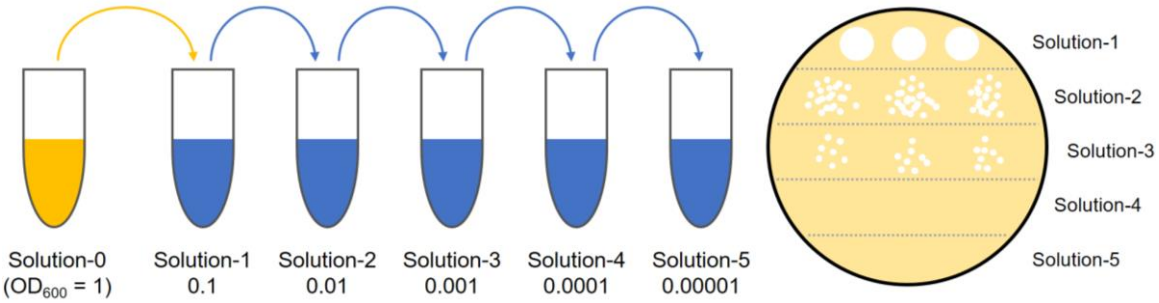


Figure 80. Illustration of serial dilution and plating for quantitative assessment of E. coli cells.

6 List of abbreviation

Å	Angstrom, 1×10^{-10} m
A	Adenosine
Abs	Absorbance
ACQ	Aggregation-caused quenching
AP	Alkaline phosphate
APS	Ammonium persulfate
ASB	Aptamer selection buffer
ATP	Adenosine triphosphate
BC	Benzopyrylium-coumarin
BHQ	Blackhole quencher
Boc	Tert-butyloxycarbonyl
BP	Bandpass
C	Cytidine
Cbl	Cobalamin
CFU	Colony-forming unit
CRISPR	Clustered regularly interspaced short palindromic repeats
CTP	Cytidine triphosphate
Cy	Cyanine
DFHBI	3,5-Difluoro-4-hydroxybenzylidine imidazolinone
DFHBI-1T	(Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1Himidazol-5(4H)-one)
DFHO	3,5-Difluoro-4-hydroxybenzylidene imidazolinone-2-oxime
DIR	Dimethyl indole red
DMF	Dimethylformamide
DMHBO	(Z)-4-(4-Hydroxy-3,5-dimethoxybenzylidene)-5-oxo-1-(4-(trimethylammonium)-phenyl)-4,5-dihydro-1H-imidazole-2-carbaldehyde oxime iodide
DMSO	Dimethyl sulfoxide
DN	Dinitroaniline
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DNB	Dinitroaniline-binding aptamer
DNFB	1-Fluoro-2,4-dinitrobenzene
dNTP	Deoxynucleoside triphosphate
ds	Double-stranded
PAGE	Polyacrylamide gel electrophoresis
PCI	Phenol/chloroform/isoamyl alcohol
PEG	Polyethylene glycol
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein

EGFR	Epidermal growth factor
eIF4A	Eukaryotic initiation factor 4A
Em	Emission
EtBr	Ethidium bromide
Ex	Excitation
FACS	Fluorescence-activated cell sorting
FISH	Fluorescent in situ hybridization
FLAP	Fluorescent light-up RNA aptamers
FP	Fluorescent protein
FRET	Fluorescence resonance energy transfer
FSC	Forward scatter
G	Guanosine
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
h	Hour
H	Height
HBC	(4-((2-Hydroxyethyl)(methyl)amino)-benzylidene)-cyanophenylacetonitrile
HBI	4-Hydroxybenzylidene imidazolinone
HCR	Hybridization chain reaction
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
Hz	Hertz
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IVT	*In vitro* transcription
JF	Janelia Fluor
K_D	Dissociation constant
L	Liter
LB	Lysogeny broth
lncRNA	Long non-coding RNA
log P	Octanol-water partition coefficient
LUMO	Lowest unoccupied molecular orbital
M	Molar
µ	Micro (10^{-6})
µIVC	microfluidic-assisted in vitro compartmentalization
m^6A	N^6-methyladenosine
MBS	MS2-binding site
MCP	MS2 coat protein
MeOH	Methanol
merFISH	Multiplexed error-robust FISH
MG	Malachite green
MGA	Malachite green-binding aptamer
MHz	MegaHertz
min	Minute

miRNA	MicroRNA
mRNA	Messenger RNA
NaOAc	Sodium acetate
NGS	Next-generation of sequencing
NHS	N-Hydroxysuccinimide
NIR	Near-infrared
NMR	Nuclear magnetic resonance
nt	Nucleotide
NTP	Nucleoside triphosphate
OD	Optical density
OTB	Oxazole and thiazole blue
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
RBP	RNA-binding protein
PCP	PP7 coat protein
PCR	Polymerase chain reaction
PDL	Poly-D-lysine
PEG	Polyethylene glycol
PFA	Paraformaldehyde
ppm	Parts per million
pre-mRNA	Pre-messenger RNA
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
ROI	RNA of interest
ROS	Reactive oxygen species
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RP	Reverse phase
RT	Reverse transcription
SDS	Sodium dodecyl sulfate
sec	Second
seqFISH	Sequential FISH
SELEX	Systematic evolution of ligand by exponential enrichment
SIM	Structured illumination microscopy
SiR	Silicon rhodamine
SiRA	Silicon rhodamine-binding aptamer
siRNA	Small interfering RNA
smFISH	Single-molecule FISH
SMLM	Single-molecule localization microscopy
S_NAr	Nucleophilic aromatic substitution
snRNA	Small nuclear RNA
SRB2	Sulforhodamine B-binding aptamer
SR-DN	Sulforhodamine-dinitroaniline
sRNA	Small RNA
ss	Single-stranded

SSC	Side scatter
STED	Stimulated emission depletion
STIC	Sequential tethered and intertwined oligonucleotides complex
T	Thymine
TBE	Tris-borate EDTA
TEMED	N,N,N′,N′-tetramethylethylendiamine
tert	Tertiary
TICT	Twisted internal charge transfer
TLC	Thin-layer chromatography
TMR	Tetramethylrhodamine
TO	Thiazole orange
Tornado	Twister-optimized RNA for durable overexpression
tRNA	Transfer RNA
U	Uridine
UTP	Uridine triphosphate
UTR	Untranslated region

7 References

1. Fu, Y.; Dominissini, D.; Rechavi, G.; He, C., Gene expression regulation mediated through reversible m(6)A RNA methylation. *Nat Rev Genet* **2014,** *15* (5), 293-306.
2. Ray, S.; Widom, J. R.; Walter, N. G., Life under the Microscope: Single-Molecule Fluorescence Highlights the RNA World. *Chem Rev* **2018,** *118* (8), 4120-4155.
3. Guerriertakada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., The Rna Moiety of Ribonuclease-P Is the Catalytic Subunit of the Enzyme. *Cell* **1983,** *35* (3), 849-857.
4. Fica, S. M.; Tuttle, N.; Novak, T.; Li, N. S.; Lu, J.; Koodathingal, P.; Dai, Q.; Staley, J. P.; Piccirilli, J. A., RNA catalyses nuclear pre-mRNA splicing. *Nature* **2013,** *503* (7475), 229-34.
5. Saldanha, R.; Mohr, G.; Belfort, M.; Lambowitz, A. M., Group-I and Group-Ii Introns. *Faseb J* **1993,** *7* (1), 15-24.
6. Waters, L. S.; Storz, G., Regulatory RNAs in Bacteria. *Cell* **2009,** *136* (4), 615-628.
7. Rana, T. M., Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Bio* **2007,** *8* (1), 23-36.
8. Pardue, M. L.; Gall, J. G., Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *Proc Natl Acad Sci U S A* **1969,** *64* (2), 600-4.
9. Femino, A.; Fay, F. S.; Fogarty, K.; Singer, R. H., Visualization of single RNA transcripts in situ. *Science* **1998,** *280* (5363), 585-590.
10. Sinnamon, J. R.; Czaplinski, K., RNA detection in situ with FISH-STICs. *Rna* **2014,** *20* (2), 260-266.
11. Choi, H. M. T.; Chang, J. Y.; Trinh, L. A.; Padilla, J. E.; Fraser, S. E.; Pierce, N. A., Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol* **2010,** *28* (11), 1208-U103.
12. Larsson, C.; Grundberg, I.; Soderberg, O.; Nilsson, M., In situ detection and genotyping of individual mRNA molecules. *Nat Methods* **2010,** *7* (5), 395-U81.
13. Chen, K. H.; Boettiger, A. N.; Moffitt, J. R.; Wang, S. Y.; Zhuang, X. W., Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **2015,** *348* (6233).
14. Lubeck, E.; Coskun, A. F.; Zhiyentayev, T.; Ahmad, M.; Cai, L., Single-cell in situ RNA profiling by sequential hybridization. *Nat Methods* **2014,** *11* (4), 360-361.
15. Pichon, X.; Lagha, M.; Mueller, F.; Bertrand, E., A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Mol Cell* **2018,** *71* (3), 468-480.
16. Morrison, L. E.; Halder, T. C.; Stols, L. M., Solution-Phase Detection of Polynucleotides Using Interacting Fluorescent Labels and Competitive Hybridization. *Anal Biochem* **1989,** *183* (2), 231-244.
17. Li, Q. Q.; Luan, G. Y.; Guo, Q. P.; Liang, J. X., A new class of homogeneous nucleic acid probes based on specific displacement hybridization. *Nucleic Acids Res* **2002,** *30* (2).
18. Cardullo, R. A.; Agrawal, S.; Flores, C.; Zamecnik, P. C.; Wolf, D. E., Detection of Nucleic-Acid Hybridization by Nonradiative Fluorescence Resonance Energy-Transfer. *P Natl Acad Sci USA* **1988,** *85* (23), 8790-8794.
19. Tyagi, S., Imaging intracellular RNA distribution and dynamics in living cells. *Nat Methods* **2009,** *6* (5), 331-338.
20. Tyagi, S.; Kramer, F. R., Molecular beacons: Probes that fluoresce upon hybridization. *Nat Biotechnol* **1996,** *14* (3), 303-308.
21. Bratu, D. P.; Cha, B. J.; Mhlanga, M. M.; Kramer, F. R.; Tyagi, S., Visualizing the distribution and transport of mRNAs in living cells. *P Natl Acad Sci USA* **2003,** *100* (23), 13308-13313.

22. Chen, A. K.; Behlke, M. A.; Tsourkas, A., Efficient cytosolic delivery of molecular beacon conjugates and flow cytometric analysis of target RNA. *Nucleic Acids Res* **2008,** *36* (12).
23. Mhlanga, M. M.; Vargas, D. Y.; Fung, C. W.; Kramer, F. R.; Tyagi, S., tRNA-linked molecular beacons for imaging mRNAs in the cytoplasm of living cells. *Nucleic Acids Res* **2005,** *33* (6), 1902-1912.
24. Santangelo, P. J.; Nix, B.; Tsourkas, A.; Bao, G., Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Res* **2004,** *32* (6).
25. Leonetti, J. P.; Mechti, N.; Degols, G.; Gagnor, C.; Lebleu, B., Intracellular-Distribution of Microinjected Antisense Oligonucleotides. *P Natl Acad Sci USA* **1991,** *88* (7), 2702-2706.
26. Wu, Y. R.; Yang, C. J.; Moroz, L. L.; Tan, W. H., Nucleic acid beacons for long-term real-time intracellular monitoring. *Anal Chem* **2008,** *80* (8), 3025-3028.
27. LeCuyer, K. A.; Behlen, L. S.; Uhlenbeck, O. C., Mutagenesis of a stacking contact in the MS2 coat protein-RNA complex. *Embo J* **1996,** *15* (24), 6847-6853.
28. Daigle, N.; Ellenberg, J., lambda(N)-GFP: An RNA reporter system for live-cell imaging. *Nat Methods* **2007,** *4* (8), 633-636.
29. Lim, F.; Downey, T. P.; Peabody, D. S., Translational repression and specific RNA binding by the coat protein of the Pseudomonas phage PP7. *J Biol Chem* **2001,** *276* (25), 22507-22513.
30. Rackham, O.; Brown, C. M., Visualization of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *Embo J* **2004,** *23* (16), 3346-3355.
31. Valencia-Burton, M.; McCullough, R. M.; Cantor, C. R.; Broude, N. E., RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat Methods* **2007,** *4* (5), 421-427.
32. Ozawa, T.; Natori, Y.; Sato, M.; Umezawa, Y., Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods* **2007,** *4* (5), 413-419.
33. Lander, E. S., The Heroes of CRISPR. *Cell* **2016,** *164* (1-2), 18-28.
34. Nelles, D. A.; Fang, M. Y.; O'Connell, M. R.; Xu, J. L.; Markmiller, S. J.; Doudna, J. A.; Yeo, G. W., Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **2016,** *165* (2), 488-496.
35. Yang, L. Z.; Wang, Y.; Li, S. Q.; Yao, R. W.; Luan, P. F.; Wu, H.; Carmichael, G. G.; Chen, L. L., Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems. *Molecular Cell* **2019,** *76* (6), 981-+.
36. Ku, T. H.; Zhang, T. T.; Luo, H.; Yen, T. M.; Chen, P. W.; Han, Y. Y.; Lo, Y. H., Nucleic Acid Aptamers: An Emerging Tool for Biotechnology and Biomedical Sensing. *Sensors-Basel* **2015,** *15* (7), 16281-16313.
37. Ellington, A. D.; Szostak, J. W., Invitro Selection of Rna Molecules That Bind Specific Ligands. *Nature* **1990,** *346* (6287), 818-822.
38. Tuerk, C.; Gold, L., Systematic Evolution of Ligands by Exponential Enrichment - Rna Ligands to Bacteriophage-T4 DNA-Polymerase. *Science* **1990,** *249* (4968), 505-510.
39. Babendure, J. R.; Adams, S. R.; Tsien, R. Y., Aptamers switch on fluorescence of triphenylmethane dyes. *J Am Chem Soc* **2003,** *125* (48), 14716-14717.
40. Liao, J. C.; Roider, J.; Jay, D. G., Chromophore-Assisted Laser Inactivation of Proteins Is Mediated by the Photogeneration of Free-Radicals. *P Natl Acad Sci USA* **1994,** *91* (7), 2659-2663.
41. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R., RNA Mimics of Green Fluorescent Protein. *Science* **2011,** *333* (6042), 642-646.
42. Strack, R. L.; Disney, M. D.; Jaffrey, S. R., A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat Methods* **2013,** *10* (12), 1219-+.
43. Filonov, G. S.; Moon, J. D.; Svensen, N.; Jaffrey, S. R., Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *J Am Chem Soc* **2014,** *136* (46), 16299-16308.

44. Li, X.; Kim, H.; Litke, J. L.; Wu, J. H.; Jaffrey, S. R., Fluorophore-Promoted RNA Folding and Photostability Enables Imaging of Single Broccoli-Tagged mRNAs in Live Mammalian Cells. *Angew Chem Int Edit* **2020,** *59* (11), 4511-4518.
45. Song, W. J.; Filonov-, G. S.; Kim, H.; Hirsch, M.; Li, X.; Moon, J. D.; Jaffrey, S. R., Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat Chem Biol* **2017,** *13* (11), 1187-+.
46. Steinmetzger, C.; Palanisamy, N.; Gore, K. R.; Hobartner, C., A Multicolor Large Stokes Shift Fluorogen-Activating RNA Aptamer with Cationic Chromophores. *Chem-Eur J* **2019,** *25* (8), 1931-1935.
47. Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocaru, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J., RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *Acs Chem Biol* **2014,** *9* (10), 2412-2420.
48. Autour, A.; Jeng, S. C. Y.; Cawte, A. D.; Abdolahzadeh, A.; Galli, A.; Panchapakesan, S. S. S.; Rueda, D.; Ryckelynck, M.; Unrau, P. J., Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat Commun* **2018,** *9*.
49. Cawte, A. D.; Unrau, P. J.; Rueda, D. S., Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat Commun* **2020,** *11* (1), 1283.
50. Chen, X. J.; Zhang, D. S.; Su, N.; Bao, B. K.; Xie, X.; Zuo, F. T.; Yang, L. P.; Wang, H.; Jiang, L.; Lin, Q. N.; Fang, M. Y.; Li, N. F.; Hua, X.; Chen, Z. D.; Bao, C. Y.; Xu, J. J.; Du, W. L.; Zhang, L. X.; Zhao, Y. Z.; Zhu, L. Y.; Loscalzo, J.; Yang, Y., Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat Biotechnol* **2019,** *37* (11), 1287-+.
51. Huang, H.; Suslov, N. B.; Li, N. S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A., A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat Chem Biol* **2014,** *10* (8), 686-U128.
52. Warner, K. D.; Chen, M. C.; Song, W. J.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferre-D'Amare, A. R., Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat Struct Mol Biol* **2014,** *21* (8), 658-663.
53. Warner, K. D.; Sjekloca, L.; Song, W. J.; Filonov, G. S.; Jaffrey, S. R.; Ferre-D'Amare, A. R., A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat Chem Biol* **2017,** *13* (11), 1195-+.
54. Trachman, R. J.; Demeshkina, N. A.; Lau, M. W. L.; Panchapakesan, S. S. S.; Jeng, S. C. Y.; Unrau, P. J.; Ferre-D'Amare, A. R., Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat Chem Biol* **2017,** *13* (7), 807-+.
55. Guo, J. U.; Bartel, D. P., RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **2016,** *353* (6306).
56. Sunbul, M.; Jaschke, A., Contact-Mediated Quenching for RNA Imaging in Bacteria with a Fluorophore-Binding Aptamer. *Angew Chem Int Edit* **2013,** *52* (50), 13401-13404.
57. Holeman, L. A.; Robinson, S. L.; Szostak, J. W.; Wilson, C., Isolation and characterization of fluorophore-binding RNA aptamers. *Fold Des* **1998,** *3* (6), 423-431.
58. Sunbul, M.; Jaschke, A., SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging. *Nucleic Acids Res* **2018,** *46* (18).
59. Sunbul, M.; Lackner, J.; Martin, A.; Englert, D.; Hacene, B.; Nienhaus, K.; Nienhaus, G. U.; Jäschke, A., RhoBAST - a rhodamine-binding aptamer for super-resolution RNA imaging. **2020**, 2020.03.12.988782.
60. Bouhedda, F.; Fam, K. T.; Collot, M.; Autour, A.; Marzi, S.; Klymchenko, A.; Ryckelynck, M., A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nat Chem Biol* **2020,** *16* (1), 69-+.
61. Murata, A.; Sato, S.; Kawazoe, Y.; Uesugi, M., Small-molecule fluorescent probes for specific RNA targets. *Chem Commun* **2011,** *47* (16), 4712-4714.

62. Sato, S.; Watanabe, M.; Katsuda, Y.; Murata, A.; Wang, D. O.; Uesugi, M., Live-Cell Imaging of Endogenous mRNAs with a Small Molecule. *Angew Chem Int Edit* **2015,** *54* (6), 1855-1858.
63. Arora, A.; Sunbul, M.; Jaschke, A., Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers. *Nucleic Acids Res* **2015,** *43* (21).
64. Braselmann, E.; Wierzba, A. J.; Polaski, J. T.; Chrominski, M.; Holmes, Z. E.; Hung, S. T.; Batan, D.; Wheeler, J. R.; Parker, R.; Jimenez, R.; Gryko, D.; Batey, R. T.; Palmer, A. E., A multicolor riboswitch-based platform for imaging of RNA in live mammalian cells. *Nat Chem Biol* **2018,** *14* (10), 964-+.
65. Wirth, R.; Gao, P.; Nienhaus, G. U.; Sunbul, M.; Jaschke, A., SiRA: A Silicon Rhodamine-Binding Aptamer for Live-Cell Super-Resolution RNA Imaging. *J Am Chem Soc* **2019,** *141* (18), 7562-7571.
66. Yerramilli, V. S.; Kim, K. H., Labeling RNAs in Live Cells Using Malachite Green Aptamer Scaffolds as Fluorescent Probes. *Acs Synth Biol* **2018,** *7* (3), 758-766.
67. Trachman, R. J.; Autour, A.; Jeng, S. C. Y.; Abdolahzadeh, A.; Andreoni, A.; Cojocaru, R.; Garipov, R.; Dolgosheina, E. V.; Knutson, J. R.; Ryckelynck, M.; Unrau, P. J.; Ferre-D'Amare, A. R., Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nat Chem Biol* **2019,** *15* (5), 472-+.
68. Trachman, R. J.; Cojocaru, R.; Wu, D.; Piszczek, G.; Ryckelynck, M.; Unrau, P. J.; Ferre-D'Amare, A. R., Structure-Guided Engineering of the Homodimeric Mango-IV Fluorescence Turn-on Aptamer Yields an RNA FRET Pair. *Structure* **2020,** *28* (7), 776-+.
69. Wang, L.; Frei, M. S.; Salim, A.; Johnsson, K., Small-Molecule Fluorescent Probes for Live-Cell Super-Resolution Microscopy. *J Am Chem Soc* **2019,** *141* (7), 2770-2781.
70. Abdelfattah, A. S.; Kawashima, T.; Singh, A.; Novak, O.; Liu, H.; Shuai, Y. C.; Huang, Y. C.; Campagnola, L.; Seeman, S. C.; Yu, J. N.; Zheng, J. H.; Grimm, J. B.; Patel, R.; Friedrich, J.; Mensh, B. D.; Paninski, L.; Macklin, J. J.; Murphy, G. J.; Podgorski, K.; Lin, B. J.; Chen, T. W.; Turner, G. C.; Liu, Z.; Koyama, M.; Svoboda, K.; Ahrens, M. B.; Lavis, L. D.; Schreiter, E. R., Bright and photostable chemigenetic indicators for extended in vivo voltage imaging. *Science* **2019,** *365* (6454), 699-+.
71. Tan, X. H.; Constantin, T. P.; Sloane, K. L.; Waggoner, A. S.; Bruchez, M. P.; Armitage, B. A., Fluoromodules Consisting of a Promiscuous RNA Aptamer and Red or Blue Fluorogenic Cyanine Dyes: Selection, Characterization, and Bioimaging. *J Am Chem Soc* **2017,** *139* (26), 9001-9009.
72. Hong, G. S.; Antaris, A. L.; Dai, H. J., Near-infrared fluorophores for biomedical imaging. *Nat Biomed Eng* **2017,** *1* (1).
73. Lukinavicius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G. Y.; Plass, T.; Mueller, V.; Reymond, L.; Correa, I. R.; Luo, Z. G.; Schultz, C.; Lemke, E. A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K., A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat Chem* **2013,** *5* (2), 132-139.
74. Yuan, L.; Lin, W. Y.; Zheng, K. B.; He, L. W.; Huang, W. M., Far-red to near infrared analyte-responsive fluorescent probes based on organic fluorophore platforms for fluorescence imaging. *Chem Soc Rev* **2013,** *42* (2), 622-661.
75. Han, K. Y.; Leslie, B. J.; Fei, J. Y.; Zhang, J. C.; Ha, T., Understanding the Photophysics of the Spinach-DFHBI RNA Aptamer-Fluorogen Complex To Improve Live-Cell RNA Imaging. *J Am Chem Soc* **2013,** *135* (50), 19033-19038.
76. Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T., Evolution of Group 14 Rhodamines as Platforms for Near-Infrared Fluorescence Probes Utilizing Photoinduced Electron Transfer. *Acs Chem Biol* **2011,** *6* (6), 600-608.

77. Wang, T.; Zhao, Q. J.; Hu, H. G.; Yu, S. C.; Liu, X.; Liu, L.; Wu, Q. Y., Spirolactonized Si-rhodamine: a novel NIR fluorophore utilized as a platform to construct Si-rhodamine-based probes. *Chem Commun* **2012,** *48* (70), 8781-8783.
78. Lukinavicius, G.; Reymond, L.; D'Este, E.; Masharina, A.; Gutfert, F.; Ta, H.; Guether, A.; Fournier, M.; Rizzo, S.; Waldmann, H.; Blaukopf, C.; Sommer, C.; Gerlich, D. W.; Arndt, H. D.; Hell, S. W.; Johnsson, K., Fluorogenic probes for live-cell imaging of the cytoskeleton. *Nat Methods* **2014,** *11* (7), 731-U168.
79. Lukinavicius, G.; Blaukopf, C.; Pershagen, E.; Schena, A.; Reymond, L.; Derivery, E.; Gonzalez-Gaitan, M.; D'Este, E.; Hell, S. W.; Gerlich, D. W.; Johnsson, K., SiR-Hoechst is a far-red DNA stain for live-cell nanoscopy. *Nat Commun* **2015,** *6*.
80. Uno, S. N.; Kamiya, M.; Yoshihara, T.; Sugawara, K.; Okabe, K.; Tarhan, M. C.; Fujita, H.; Funatsu, T.; Okada, Y.; Tobita, S.; Urano, Y., A spontaneously blinking fluorophore based on intramolecular spirocyclization for live-cell super-resolution imaging. *Nat Chem* **2014,** *6* (8), 681-689.
81. Grimm, J. B.; Muthusamy, A. K.; Liang, Y. J.; Brown, T. A.; Lemon, W. C.; Patel, R.; Lu, R. W.; Macklin, J. J.; Keller, P. J.; Ji, N.; Lavis, L. D., A general method to fine-tune fluorophores for live-cell and in vivo imaging. *Nat Methods* **2017,** *14* (10), 987-+.
82. Wang, L.; Tran, M.; D'Este, E.; Roberti, J.; Koch, B.; Xue, L.; Johnsson, K., A general strategy to develop cell permeable and fluorogenic probes for multicolour nanoscopy. *Nat Chem* **2020,** *12* (2), 165-+.
83. Czerney, P.; Graness, G.; Birckner, E.; Vollmer, F.; Rettig, W., Molecular Engineering of Cyanine-Type Fluorescent and Laser-Dyes. *J Photoch Photobio A* **1995,** *89* (1), 31-36.
84. Liu, J.; Sun, Y. Q.; Wang, P.; Zhang, J. Y.; Guo, W., Construction of NIR and ratiometric fluorescent probe for Hg2+ based on a rhodamine-inspired dye platform. *Analyst* **2013,** *138* (9), 2654-2660.
85. Lee, M. H.; Kim, J. S.; Sessler, J. L., Small molecule-based ratiometric fluorescence probes for cations, anions, and biomolecules. *Chem Soc Rev* **2015,** *44* (13), 4185-4191.
86. Palmer, A. E.; Qin, Y.; Park, J. G.; McCombs, J. E., Design and application of genetically encoded biosensors. *Trends Biotechnol* **2011,** *29* (3), 144-152.
87. Lv, H. M.; Yang, X. F.; Zhong, Y. G.; Guo, Y.; Li, Z.; Li, H., Native Chemical Ligation Combined with Spirocyclization of Benzopyrylium Dyes for the Ratiometric and Selective Fluorescence Detection of Cysteine and Homocysteine. *Anal Chem* **2014,** *86* (3), 1800-1807.
88. Ding, S. S.; Zhang, Q.; Xue, S. H.; Feng, G. Q., Real-time detection of hypochlorite in tap water and biological samples by a colorimetric, ratiometric and near-infrared fluorescent turn-on probe. *Analyst* **2015,** *140* (13), 4687-4693.
89. Duan, Y. W.; Yang, X. F.; Zhong, Y. G.; Guo, Y.; Li, Z.; Li, H., A ratiometric fluorescent probe for gasotransmitter hydrogen sulfide based on a coumarin-benzopyrylium platform. *Anal Chim Acta* **2015,** *859*, 59-65.
90. Dai, X.; Wang, Z. Y.; Du, Z. F.; Miao, J. Y.; Zhao, B. X., A simple but effective near-infrared ratiometric fluorescent probe for hydrazine and its application in bioimaging. *Sensor Actuat B-Chem* **2016,** *232*, 369-374.
91. Dong, B. L.; Song, X. Z.; Kong, X. Q.; Wang, C.; Tang, Y. H.; Liu, Y.; Lin, W. Y., Simultaneous Near-Infrared and Two-Photon In Vivo Imaging of H2O2 Using a Ratiometric Fluorescent Probe based on the Unique Oxidative Rearrangement of Oxonium. *Adv Mater* **2016,** *28* (39), 8755-8759.
92. Flores-Cruz, R.; Lopez-Arteaga, R.; Ramirez-Vidal, L.; Lopez-Casillas, F.; Jimenez-Sanchez, A., Unravelling the modus-operandi of chromenylium-cyanine fluorescent probes: a case study. *Phys Chem Chem Phys* **2019,** *21* (28), 15779-15786.
93. Grimm, J. B.; Tkachuk, A. N.; Xie, L. Q.; Choi, H.; Mohar, B.; Falco, N.; Schaefer, K.; Patel, R.; Zheng, Q. S.; Liu, Z.; Lippincott-Schwartz, J.; Brown, T. A.; Lavis, L. D., A general method to optimize and functionalize red-shifted rhodamine dyes. *Nat Methods* **2020,** *17* (8), 815-+.

94. Wang, L. L.; Du, W.; Hu, Z. J.; Uvdal, K.; Li, L.; Huang, W., Hybrid Rhodamine Fluorophores in the Visible/NIR Region for Biological Imaging. *Angew Chem Int Edit* **2019,** *58* (40), 14026-14043.
95. Butkevich, A. N.; Mitronova, G. Y.; Sidenstein, S. C.; Klocke, J. L.; Kamin, D.; Meineke, D. N. H.; D'Este, E.; Kraemer, P. T.; Danzl, J. G.; Belov, V. N.; Hell, S. W., Fluorescent Rhodamines and Fluorogenic Carbopyronines for Super-Resolution STED Microscopy in Living Cells. *Angew Chem Int Edit* **2016,** *55* (10), 3290-3294.
96. Sakabe, M.; Asanuma, D.; Kamiya, M.; Iwatate, R. J.; Hanaoka, K.; Terai, T.; Nagano, T.; Urano, Y., Rational Design of Highly Sensitive Fluorescence Probes for Protease and Glycosidase Based on Precisely Controlled Spirocyclization. *J Am Chem Soc* **2013,** *135* (1), 409-414.
97. Chiba, M.; Ichikawa, Y.; Kamiya, M.; Komatsu, T.; Ueno, T.; Hanaoka, K.; Nagano, T.; Lange, N.; Urano, Y., An Activatable Photosensitizer Targeted to gamma-Glutamyltranspeptidase. *Angew Chem Int Edit* **2017,** *56* (35), 10418-10422.
98. Iwatate, R. J.; Kamiya, M.; Urano, Y., Asymmetric Rhodamine-Based Fluorescent Probe for Multicolour In Vivo Imaging. *Chem-Eur J* **2016,** *22* (5), 1696-1703.
99. Li, Z.; Wu, S. Q.; Han, J. H.; Han, S. F., Imaging of intracellular acidic compartments with a sensitive rhodamine based fluorogenic pH sensor. *Analyst* **2011,** *136* (18), 3698-3706.
100. Dujols, V.; Ford, F.; Czarnik, A. W., A long-wavelength fluorescent chemodosimeter selective for Cu(II) ion in water. *J Am Chem Soc* **1997,** *119* (31), 7386-7387.
101. Kenmoku, S.; Urano, Y.; Kojima, H.; Nagano, T., Development of a highly specific rhodamine-based fluorescence probe for hypochlorous acid and its application to real-time imaging of phagocytosis. *J Am Chem Soc* **2007,** *129* (23), 7313-7318.
102. Chao, J. A.; Lionnet, T., Imaging the Life and Death of mRNAs in Single Cells. *Csh Perspect Biol* **2018,** *10* (12).
103. Komarova, N.; Kuznetsov, A., Inside the Black Box: What Makes SELEX Better? *Molecules* **2019,** *24* (19).
104. Green, N. M., Avidin. *Adv Protein Chem* **1975,** *29*, 85-133.
105. Davis, J. H.; Szostak, J. W., Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *P Natl Acad Sci USA* **2002,** *99* (18), 11616-11621.
106. Shedlovskiy, D.; Shcherbik, N.; Pestov, D. G., One-step hot formamide extraction of RNA from Saccharomyces cerevisiae. *Rna Biol* **2017,** *14* (12), 1722-1726.
107. Grimm, J. B.; English, B. P.; Chen, J. J.; Slaughter, J. P.; Zhang, Z. J.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; Lionnet, T.; Lavis, L. D., A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat Methods* **2015,** *12* (3), 244-+.
108. Nazarenko, I.; Pires, R.; Lowe, B.; Obaidy, M.; Rashtchian, A., Effect of primary and secondary structure of oligodeoxyribonucleotides on the fluorescent properties of conjugated dyes. *Nucleic Acids Res* **2002,** *30* (9), 2089-2095.
109. Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M., Nucleobase-specific quenching of fluorescent dyes .1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *J Phys Chem-Us* **1996,** *100* (13), 5541-5553.
110. Hwang, G. T., Single-Labeled Oligonucleotides Showing Fluorescence Changes Upon Hybridization with Target Nucleic Acids. *Molecules* **2018,** *23* (1).
111. Corpet, F., Multiple Sequence Alignment with Hierarchical-Clustering. *Nucleic Acids Res* **1988,** *16* (22), 10881-10890.
112. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., Selection of Single-Stranded-DNA Molecules That Bind and Inhibit Human Thrombin. *Nature* **1992,** *355* (6360), 564-566.
113. Hesselberth, J. R.; Miller, D.; Robertus, J.; Ellington, A. D., In vitro selection of RNA molecules that inhibit the activity of ricin A-chain. *J Biol Chem* **2000,** *275* (7), 4937-4942.
114. Trachman, R. J., 3rd; Abdolahzadeh, A.; Andreoni, A.; Cojocaru, R.; Knutson, J. R.; Ryckelynck, M.; Unrau, P. J.; Ferre-D'Amare, A. R., Crystal Structures of the Mango-II RNA

Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness. *Biochemistry-Us* **2018,** *57* (26), 3544-3548.
115. Wang, L.; Hiblot, J.; Popp, C.; Xue, L.; Johnsson, K., Environmentally Sensitive Color-Shifting Fluorophores for Bioimaging. *Angew Chem Int Ed Engl* **2020**.
116. Mitronova, G. Y.; Belov, V. N.; Bossi, M. L.; Wurm, C. A.; Meyer, L.; Medda, R.; Moneron, G.; Bretschneider, S.; Eggeling, C.; Jakobs, S.; Hell, S. W., New fluorinated rhodamines for optical microscopy and nanoscopy. *Chemistry* **2010,** *16* (15), 4477-88.
117. Sarkar, M.; Poddar, S., Studies on the interaction of surfactants with cationic dye by absorption spectroscopy. *J Colloid Interf Sci* **2000,** *221* (2), 181-185.
118. Misra, V. K.; Draper, D. E., The linkage between magnesium binding and RNA folding. *J Mol Biol* **2002,** *317* (4), 507-521.
119. Draper, D. E., A guide to ions and RNA structure. *Rna* **2004,** *10* (3), 335-343.
120. Walmsley, J. A.; Burnett, J. F., A new model for the K+-induced macromolecular structure of guanosine 5 '-monophosphate in solution. *Biochemistry-Us* **1999,** *38* (42), 14063-14068.
121. Ueyama, H.; Takagi, M.; Takenaka, S., A novel potassium sensing in aqueous media with a synthetic oligonucleotide derivative. Fluorescence resonance energy transfer associated with guanine quartet-potassium ion complex formation. *J Am Chem Soc* **2002,** *124* (48), 14286-14287.
122. Pogliano, J.; Ho, T. Q.; Zhong, Z.; Helinski, D. R., Multicopy plasmids are clustered and localized in Escherichia coli. *Proc Natl Acad Sci U S A* **2001,** *98* (8), 4486-91.
123. Ponchon, L.; Dardel, F., Recombinant RNA technology: the tRNA scaffold. *Nat Methods* **2007,** *4* (7), 571-576.
124. Balleza, E.; Lopez-Bojorquez, L. N.; Martinez-Antonio, A.; Resendis-Antonio, O.; Lozada-Chavez, I.; Balderas-Martinez, Y. I.; Encarnacion, S.; Collado-Vides, J., Regulation by transcription factors in bacteria: beyond description. *Fems Microbiol Rev* **2009,** *33* (1), 133-151.
125. Paige, J. S.; Nguyen-Duc, T.; Song, W.; Jaffrey, S. R., Fluorescence imaging of cellular metabolites with RNA. *Science* **2012,** *335* (6073), 1194.
126. Wu, R.; Karunanayake Mudiyanselage, A.; Shafiei, F.; Zhao, B.; Bagheri, Y.; Yu, Q.; McAuliffe, K.; Ren, K.; You, M., Genetically Encoded Ratiometric RNA-Based Sensors for Quantitative Imaging of Small Molecules in Living Cells. *Angew Chem Int Ed Engl* **2019,** *58* (50), 18271-18275.
127. Litke, J. L.; Jaffrey, S. R., Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat Biotechnol* **2019,** *37* (6), 667-675.
128. Wittmann, A.; Suess, B., Selection of tetracycline inducible self-cleaving ribozymes as synthetic devices for gene regulation in yeast. *Mol Biosyst* **2011,** *7* (8), 2419-2427.
129. Zhou, T.; Damsky, W.; Weizman, O. E.; McGeary, M. K.; Hartmann, K. P.; Rosen, C. E.; Fischer, S.; Jackson, R.; Flavell, R. A.; Wang, J.; Sanmamed, M. F.; Bosenberg, M. W.; Ring, A. M., IL-18BP is a secreted immune checkpoint and barrier to IL-18 immunotherapy. *Nature* **2020,** *583* (7817), 609-+.
130. Ouisse, L. H.; Gautreau-Rolland, L.; Devilder, M. C.; Osborn, M.; Moyon, M.; Visentin, J.; Halary, F.; Bruggemann, M.; Buelow, R.; Anegon, I.; Saulquin, X., Antigen-specific single B cell sorting and expression-cloning from immunoglobulin humanized rats: a rapid and versatile method for the generation of high affinity and discriminative human monoclonal antibodies. *Bmc Biotechnol* **2017,** *17*.
131. Ambriz-Avina, V.; Contreras-Garduno, J. A.; Pedraza-Reyes, M., Applications of flow cytometry to characterize bacterial physiological responses. *Biomed Res Int* **2014,** *2014*, 461941.
132. Muller, S.; Nebe-von-Caron, G., Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. *Fems Microbiol Rev* **2010,** *34* (4), 554-87.
133. Dvorak, P.; Chrast, L.; Nikel, P. I.; Fedr, R.; Soucek, K.; Sedlackova, M.; Chaloupkova, R.; de Lorenzo, V.; Prokop, Z.; Damborsky, J., Exacerbation of substrate toxicity by IPTG in

Escherichia coli BL21(DE3) carrying a synthetic metabolic pathway. *Microb Cell Fact* **2015,** *14*, 201.
134. Garcia-Fraga, B.; da Silva, A. F.; Lopez-Seijas, J.; Sieiro, C., Optimized expression conditions for enhancing production of two recombinant chitinolytic enzymes from different prokaryote domains. *Bioproc Biosyst Eng* **2015,** *38* (12), 2477-2486.
135. Bacterial Systems. In *Production of Membrane Proteins*, pp 11-35.
136. Hemadou, A.; Laroche-Traineau, J.; Antoine, S.; Mondon, P.; Fontayne, A.; Le Priol, Y.; Claverol, S.; Sanchez, S.; Cerutti, M.; Ottones, F.; Clofent-Sanchez, G.; Jacobin-Valat, M. J., An innovative flow cytometry method to screen human scFv-phages selected by in vivo phage-display in an animal model of atherosclerosis. *Sci Rep-Uk* **2018,** *8*.
137. Song, J. N.; Matthews, A. Y.; Reboul, C. F.; Kaiserman, D.; Pike, R. N.; Bird, P. I.; Whisstock, J. C., Predicting Serpin/Protease Interactions. *Method Enzymol* **2011,** *501*, 237-273.
138. Wirth, R.; Jäschke, A., *Development of a Genetically Encoded Fluorescent Light-up Aptamer for Live-cell Super-resolution RNA Imaging*. Ruperto Carola University Heidelberg: 2020.
139. Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W., HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991,** *67* (3), 529-36.
140. Filonov, G. S.; Kam, C. W.; Song, W. J.; Jaffrey, S. R., In-Gel Imaging of RNA Processing Using Broccoli Reveals Optimal Aptamer Expression Strategies. *Chem Biol* **2015,** *22* (5), 649-660.
141. Romero, I. G.; Pai, A. A.; Tung, J.; Gilad, Y., RNA-seq: impact of RNA degradation on transcript quantification. *Bmc Biol* **2014,** *12*.
142. Zhao, Y. Z.; Jin, J.; Hu, Q. X.; Zhou, H. M.; Yi, J.; Yu, Z. H.; Xu, L.; Wang, X.; Yang, Y.; Loscalzo, J., Genetically Encoded Fluorescent Sensors for Intracellular NADH Detection. *Cell Metab* **2011,** *14* (4), 555-566.
143. Tao, R. K.; Zhao, Y. Z.; Chu, H. Y.; Wang, A. X.; Zhu, J. H.; Chen, X. J.; Zou, Y. J.; Shi, M.; Liu, R. M.; Su, N.; Du, J. L.; Zhou, H. M.; Zhu, L. Y.; Qian, X. H.; Liu, H. Y.; Loscalzo, J.; Yang, Y., Genetically encoded fluorescent sensors reveal dynamic regulation of NADPH metabolism. *Nat Methods* **2017,** *14* (7), 720-+.
144. Mank, M.; Griesbeck, O., Genetically encoded calcium indicators. *Chem Rev* **2008,** *108* (5), 1550-1564.
145. Wilms, C. D.; Hausser, M., Lighting up neural networks using a new generation of genetically encoded calcium sensors. *Nat Methods* **2009,** *6* (12), 871-872.
146. Song, W. J.; Strack, R. L.; Jaffrey, S. R., Imaging bacterial protein expression using genetically encoded RNA sensors. *Nat Methods* **2013,** *10* (9), 873-+.
147. Kellenberger, C. A.; Wilson, S. C.; Sales-Lee, J.; Hammond, M. C., RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP. *J Am Chem Soc* **2013,** *135* (13), 4906-9.
148. Ratz, M.; Testa, I.; Hell, S. W.; Jakobs, S., CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci Rep-Uk* **2015,** *5*.
149. Houseley, J.; Tollervey, D., The Many Pathways of RNA Degradation. *Cell* **2009,** *136* (4), 763-776.
150. Critchfield, F. E.; Gibson, J. A.; Hall, J. L., Dielectric Constant for the Dioxane Water System from 20 to 35-Degrees. *J Am Chem Soc* **1953,** *75* (8), 1991-1992.
151. Wurth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-Genger, U., Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nat Protoc* **2013,** *8* (8), 1535-50.

8 Appendix

Library	Rank	Sequence
L2-R3-L	1	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTCCCTCAAGGGAACGCTT
	2	GACGTTTGAAAACGTCTAACACGAGTCTGCTTCGGCAGAGTCTGACGGTATCCCGGCGGATGTT
	3	AGCTGGCGTGGTGTATAGTCTCCTGGCTGCTTCGACAGCTGTTTAAATCGATCTGGCGGACATT
	4	ACACTTTGAAAAGTGTCACTGCTATGCTGCTTCGGCAGCATCTCACGGTATCCCGGCGGACGA
	5	CGCCCCCACCGGGTTTGAAAACCTGGCTGCTTCGGCAGTTGTATCCTTTGGGGCTCGGCAATTC
	6	ACTCAGATGGTTTTGAAAAGCCATGACTGCTTCGGCAGTTGATATCTCCTGAACTCCTGATGGA
	7	AGCGTCGGTGGTGTATAGTCGGCGTCCTGCAGCAGTATTGGCTCGGTTCGGCCCCCTCATC
	8	GTGGTGTATAGTTCCTGCGATGGCATCTGCTTCGGCAGATATACTGGGATCCGTGACGATCATC
	9	TACGTACGGACTGAGCAAAGAGCTAACTACTTCGGCAGACCCTGCCCGGACGGCGACCCCGGTCT
	10	TTCTCGTGAGTGCCCGCGCTCGCTGGTGTCGCGGATCAGTTTTTGAATCCACTCAAGGTTAACGATA
L1-R3-L	1	AGTAGGATCAAGGTGCTGTAGGTGTTCTGCTTCGGCAGAGGTGTCCTACTGATAGCTTCCGCCC
	2	AGTCGCCCGGGTGTATAGGTATCGACTGCTTCGGCAGATAGCTCAAGATGGACCGGTACCATT
	3	Same as L2-R3-L-8
	4	Same as L2-R3-L-3
	5	Same as L2-R3-L-7
	6	Same as L2-R3-L-5
	7	Same as L2-R3-L-1
	8	GTGATGTATAGCCCCAGTGAACTATCCTGCTTCGGCAGACATATGCTCCGGTCCGCCGGGCATC
	9	Same as L2-R3-L-9
	10	GTTGCTGTTTGCGGTACACGCCGTGCTGGCGGCCTCAATGTCGACCGTAATGTATCTCTTTGGTGTGGTG

Appendix Table 1. List of top 10 RNA sequences (by frequency according to NGS data) from L1-R3-L and L2-R3-L. The primer binding sequences at the 5′ and 3′ ends are not shown.

Library	RNA	Sequence	Freq
L1-R3-L	A03	AGTCGCCCGGGTGTATAGGTATCGACTGCTTCGGCAGATAGCTCAAGATGGACCGGTACCATT	3
	A04_1	TTCTCGTGAGTGCCCGCGCTCGCTGGTGTCGCGGATCAGTTCTTGAATCCACTCAAGGTTAACGATA	1
	A04_2	GTGGTGTATAGTTCCTGCGATGGCATCTGCTTCGGCAGATATACTGGGATCCGTGACGATCATC	1
	A04_3	AGCGTCGGTGGTGTATAGTCGGCGTCCTGCGGCAGTATTCGCTCGGTTCGGCCCCCTCATC	3
	A04_4	TCATCCAGCGAGCGGAAATCACCCTCATCCCCACGGTCGGCGCGAGGCGGATCGCA	1
	B01	TGCCGCGAGGCGGTAAAACGCATGGCGCTTAAAAATGGTGGCAGTGGCGGCGCTATCGTCAATGTCTCTTCGGT	1
	C02_2	TGGGAAAAGTACGGCCTGAGGCCCGACCAGTGGGCCGACTACCGGGCCCTGACTG	2
	C03	GGAACACCATTAAGGGTATGATGGAGGGGCGGGGGTTATCTGCGGCGGGTTTG	1
	D02	CGCCCCCACCGGGTTTGAAAACCTGGCTGCTTCGGCAGTTGTATCCTTTGGGGCTCGGCAATTC	1
	D04	GTGATGTATAGCCCCAGTGAACTATCCTGCTTCGGCAGACATATGCTCCGGTCCGCCGGGCATC	2
	F04	CCGCGATGTAGAGTTTACTTTTCNAAAATTTATGACCGGANGCGTGCCGCAATTTCNTCNGGTCTANAAA	1
	H01	NGATTGACNCGGTGGCGAGGATGGAGGCGCGCTAGGTTACGGGCCTGCAGGATTG	1
	H03_1	TTTGGTAGCGTTTGAAAACGTATACCCTGCTTCGGCAGTATCCTATCGGAACCAAAGTCCGTGAG	1
	H04	AGTAGGATCAAGGTGCTGTAGGTGTTCTGCTTCGGCAGAGGAGTCCTACTGATAGCTTCCGCCC	1
L1-R3	A03_2	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTACCTCAAGGGAACGCTT	2
	E04_1	TTCTCGTGAGTGCCCGCGCTCGCTGGTGTCGCGGATCAGTTTTTGAATCCACTCAAGGTTAACGATA	1
L2-R3-L	B06	CCGAAGTTGATGGTCTTGGCCGCCCGGCGCATCAGGGGGTCCACGGCCT	1
	C05_2	NTGTTTANCTGCAGAGCANTGAGGNNNCANGGNCACCTCGGTCNGATNCACTCNAATACGACTCACTATA	1
	C07	CTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACC	1
	C08_2	CGCCCCCACCGGGTTTGAAAACCTGGCTGCTTCGGCAGTTGTATCCTTTAGGGCTCGGCAACTC	1
	D08_2	TNTAACACGACTCANTATAGGAGCTCANCCTTCACNGCAGTAAAACANTCNCTN	1
	E07	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTCCCTCAAGGGAACGCTT	1
	G06_1	AAGATGTGGACCATTTAACTTGTAGACTGCCTCGGCAGGCGGCTGTTCCCTCAAGGGAACGCTT	2
	C05	ATTCAGATGGTTTTGAAAAGCCATGACTGCTTCGGCAGTTGATATCTCCTGAACTCCTGATGGA	2

	G06_2	GCCAAAGGAGCCTCTCCTCCCCCTCAAGCCTCCCCCACAGGTTGGCGA	1
L2-R3	A06	GCCATACTGCTCGCGCGCCTGCGGTTCCGTTAAACCAACAGTACCAATCGGCGGAT	2
	B06	TACTCGGATCGGCCTTGAAAGACCGTCTGCTTCGGCAGGGTATCCCTATGAGCGCAGGTACCTA	1
	B07	TCAAAATCATTACGCGCTTCATAGCGTGGCGGCACCACTTGCTTCATC	1
	C05_2	ACACTTTGAAAAGTGTCCCTGCTATGCTGCTTCGGCAGCATCTCACGGTATCCCGGCGGACGA	2
	C05_3	GTTCTATCGGTAATGCAGTTTGAAAACTGCTTCGCAGTCGTAGCACGGCCCAACGCTTGCTTC	2
	C05_4	TACGTAYGGACTGAGCAAAKAGCTAACTACTTCGGCAGACCCCTGCCGGACGSCGACCCCRGTCT	1
	C07	ACACTTTGAAAAGTGTCACTGCTGTGCTGCTTCGGCAGCATCTCACGGTATCCCGGCGGTCGA	1
	C08_1	ACACTTTGAAAAGTGTCACTGCTATGCTGCTTCGGCAGCATCTCACGGTATCCCGGCGGATGTT	1
	C08_2	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTCCCCCAAGGGAACGCTT	1
	F07_1	AGCTGGCGTGGTGTATAGTCTCCTGGCTGCTTCGACAGCTGTTTAAATCGATCTGGCGGACATT	1
	F07_2	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCAGCTGTTCCCTCAAGGGAACGCTT	1
	F07_3	ACACTTTGAAAAGTGTCACTGCTATGCTGCTTCGGTAGCATCTCACGGTATCCCGGCGGACGA	1
	G07	TGAACTTTGGTATGTAAACCGACCTTCTGCTTCGGCAGTACTCCACGACCTGTCAGTAGGTAG	1
	H08_1	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTCCCTCAAGGGAACGCTT	1
	G07	TACTCGGATCGGTCTTGAAAGACCGTCTGCCTCGGCAGGGNATCCCTANNAGCGCAGGTACNTA	1
L3-R3-L	A09_2	GGTCATCGGGTTTGAAAACCCGGCGGCTTCGGCAGTTTTATGCTTTGGCC	1
	A12	TGCAACCGGGTTTGGTTACCTTGCTGCTGCGGGAGTCTCATCCCATGGAA	1
	C12	CTCCACCGGGGTTGAAAAACTTGCTACGTCGGCAGTTCTATGCTTTGGCA	1
	D09_2	ATTCAGATGGTTTTGAAAAGCCATGACTGCTTCGGCAGTTGATATCTCCTGAACTGATGGA	1
	F09	TACGTACGGACTGAGCAAAGAGCTAACTACTTCGGCAGACCCTGCCGGACGGCGACTCCGGTCT	1
	E12	GGCCGCCGGGTATGAAAACCTGGTTGCTTTGGCAGTTGTATCCTCTAGCC	1
	G10	GGCGACCTGGTTTGATAACCTTGCTGCTTCGGCAGTTGTATCCTCTGGCT	1
	H09_1	GGCCACCGGGTTTGAAGACCTGGGTGCTTCGGAAGTTTTATCCATTGGCA	1
	H09_2	GCGGTAGATATTGCTGTTGATCCGATTGAAGGCACGCGCAT	1

	C09_1	TCAACATTTTGCACCGGAATCGGCTCGTACTTGTGTACGATGTTGCCAACATCATCTTCTTTCCCTTCCTG	1
	C09_2	AAGATGTGGACCATTTAACTTGTAGACTGCTTCGGCAGGCGGCTGTTCCCTCAAGGGAACGCTN	1
L3-R3	A09	GGTGGCCACCGGGATTGAAAACCTGGTTGCTTCGGCAGTAGTATCCATGGGCT	1
	A12	GGCCCCCAAGACCTGGGTACTTCGGCAGTTGTCTCCTTTGGTC	1
	B11	GTCAACCACGTTTAAAAACCTGGCTGCTTCGGCCGTTGTATACTTTGGCC	1
	C09_1	CAGAGGGTGAGGTNCGCGAAGTAGTGGTCGATTAAACAGAGCGC	1
	C09_2	GGCCTCAGGGCTTGAGAACCTGGGTGCTGCGGCAGTTGGTTCCTCTGTCC	1
	E10	GTCGATGCGGGTCCAGGTCGGTAACACGTCACGGTTTT	1
	E12_1	GGCCACCAGGTTTGCAAACCTGGAGTCTTGGGCAGTTGTATTCTTTGACC	1
	E12_2	TGCCAGCGGGTTTGATAACCTGGCTGCTTCGGCAGTTTTATCCTTTTGCC	1
	G11	GCGGTAGATATTGCTGTTGATCCGATTGAAGGCACGCGCAT	1
	H11_1	AGTGGCCACTGGATTTGAAAACCTGGCTGCTTCACCAGTTCTATCGTTTGGCC	1
	A12_1	GGCCGCCGCGTTAGCAAACCTGGCCGGTTCGGCAGTTTTATCGTTTGGCA	1
	A12_2	GGCCACCGGGTTTGAATACCCGGCTTCTTCGGCAGTTGTATCCGTTGGCG	1
	B12	TGCAWCCGGGTTTGAAAACCTGGCCGCTTCGGTAGTTGTATCCTTTGGCC	1
	F09	AGCCACTGGGTTTGTAAAACTGGCTGCTTCGTCTGTTGTATCCTACGTCC	1
	G09_1	GGCGGCCTCCGGTTTTGAATACCTGTCTGCTTCAGCAGTAGTATCCTCTGTCC	1
	G09_2	NGCACCGGGTTTGAAAACCTCGATGNTTCGGGAGTTGTATCCTTTGGCC	1
	G10	AATACAGCCATTAGAAAAACTGCCTATCGATTGTGCGCTACCCGGCACAGGTTGGGTTATTTTTTGCCAT	1
	H11_2	TACGTACGGACTGAGCAAAGAGCTAACTACTTCGGCAGACCCTGCCGGACGGCGACCCCGGTCT	1

Appendix Table 2. Sanger sequencing results of the selected clones form the output DNA pools of L1-R3-L, L1-R3, L2-R3-L, L2-R3, L3-R3-L, and L3-R3, respectively. The primer binding sequences at the 5′ and 3′ ends of each clone are not shown.

Eidesstattliche Erklärung

Eidesstattliche Versicherung gemäß §8 der Promotionsordnung für die Naturwissenschaftlich-Mathematische Gesamtfakultät der Universität Heidelberg

1. Bei der eingereichten Dissertation zu dem Thema „ Development of Genetically Encoded Near-Infrared Fluorescent Light-up Aptamers for Live-Cell RNA imaging " handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer anderen Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

…….. ……..

Ort und Datum Unterschrift