Supporting Information

Genetically Encoded Catalytic Hairpin Assembly for Sensitive RNA Imaging in Live Cells

Aruni P.K.K. Karunanayake Mudiyanselage, Qikun Yu, Mark A. Leon-Duque, Bin Zhao, Rigumula Wu, and Mingxu You*

Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003, USA

* mingxuyou@chem.umass.edu

This file includes:

Materials and Methods

Reagents: All chemicals were purchased from Sigma or Fisher Scientific unless otherwise noted. All the RNA structures and circuits were designed and optimized using the NUPACK and Mfold software. DNA oligonucleotides were synthesized and purified by W. M. Keck Oligonucleotide Synthesis Facility (Yale University School of Medicine) and Integrated DNA Technologies (Coralville, IA). Oligonucleotides were dissolved at 100 µM concentration in 10 mM Tris–HCl, 0.1 mM EDTA at pH=7.5 and stored at -20°C. Double stranded DNA template and non-template strands for *in vitro* transcription was prepared by PCR amplification using an Eppendorf Mastercycler. PCR product was gel purified and cleaned using a QIAquick PCR purification kit (Qiagen, Germantown, MD). All the concentrations of nucleic acids were measured using a NanoDrop One UV-Vis spectrophotometer. All the RNAs for *in vitro* test was transcribed using a HiScribe™ T7 high yield RNA synthesis kit (New England BioLabs, Ipswich, MA) and the desired product was gel purified with 10% denaturing PAGE using established protocols. All the RNA molecules were prepared into aliquots and stored at -20°C for immediate usage or at -80°C for long-term storage. RNA from cells were isolated according to Invitrogen user guide with TRIzole reagent.

Fluorescence assay: All the *in vitro* fluorescence measurements were conducted in a PTI fluorimeter (Horiba, New Jersey, NJ). RNA molecules were pre-heated at 95°C for 3 min and slowly cooled down to 25°C at the rate of -3°C/min. Fluorescence assays and RNA assembly reactions were conducted in a buffer consisting of 10 mM Tris, 5 mM $MgCl_2$, 100 mM KCl, and 10 mM NaCl at pH=7.5. For the buffer condition optimization, we used a phosphate buffer consists of 50 mM K_2HPO_4, 100 mM KCl, 5 mM $MgCl_2$ and 10 mM NaCl at pH= 7.5, and a HEPES buffer consists of 20 mM HEPES, 5 mM $MgCl_2$, 100 mM KCl and 10 mM NaCl at pH= 7.5. 250 nM H1 and H2 was used for these measurement, with the addition of 0.5 – 250 nM concentration of target RNA unless stated otherwise. All the reactions were initiated by adding H1 to the mixture and the samples have been incubated for 2 h before taking the measurements. The 500 – 550 nm emission spectra were collected by exciting at 480 nm. The kinetic assays were conducted by exciting at 480 nm and collecting the fluorescence data at 503 nm. All the data were plotted using the Origin software.

Vector construction: H1 and H2 hairpins were cloned into a pETDuet vector. The vector was first digested with NdeI and PacI restriction enzymes (New England BioLabs). After gel purification, the digested vector was ligated with a similarly digested H2 insert using T4 DNA ligase (New England BioLabs). The ligated product was transform into BL21 (DE3)* cells (New England BioLabs) with ampicillin resistance. H1 sequence was cloned into the pETDuet vector with or without H2 insert using a Gibson Assembly® cloning kit (New England BioLabs). The catalyst, molecular beacon and theophylline-regulated HHR construct was cloned, respectively, into a pCDFDuet vector system (EMD Millipore, Burlington, MA). After digestion with EcoNI and HindIII restriction enzymes (New England BioLabs) and gel purification, the digested vector was ligated with a similarly digested insert. The ligated product was transformed into BL21 (DE3)* cells with streptomycin resistance. All these plasmids were isolated and confirmed by Sanger sequencing at Genewiz, NJ.

Cellular imaging and data analysis: Cellular imaging was performed according to a previously established protocol [1]. The BL21 (DE3)* cells were grown in LB media at 37°C until the optical

density at 600 nm reaching 0.4-0.5, and then 1 mM IPTG was added for a 2 h induction before prepared for the confocal imaging. In the SgrS detection assay, cells were IPTG induced for 1.5 h and then treated with glucose (1 g/L and 20 g/L) for an additional 30 min before imaging. For the theophylline-regulated RNA imaging, cells were induced by IPTG for 1 h before adding different concentrations of theophylline for an additional 30 min or 1 h before imaging. All fluorescence images were collected with a NiS-Elements AR software using a Yokogawa spinning disk confocal on a Nikon Eclipse-TI inverted microscope. Broccoli and RNA sensors were excited with a 488 nm laser with a 60× oil immersion objective. Data analysis was performed with the NiS-Elements AR Analysis software. Data calculation and fitting was done using the Origin software. All the gel analysis was done using the ImageJ software.

Table S1: Corresponding DNA sequences of hairpins and target catalysts used in this project. Sequences for the Broccoli aptamer are shown in italic. Sequences that are modified from the initial design 1 are underlined. The different nucleic acids between D3 and D4 are bolded.

Design	Sequence
D1	**H1**:GAGCGATACTGTTATTAGATGTCGCCATGTCTTGCGACATCTAATAACAGCGAGACAA*CAGAGAG ACGGTCGGGTCCACAGTTTCC* **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTCTCGAGATGTCGCAAGACATGGCGACAT CTAATAACAGCCATGTCTTG **Target**: GACATCTAATAACAGTATCGCT
D2	**H1**:GAGCGATACTGTTGGTAGATGTCGCCATGTCTTGCGACATCTACCAACAGCGAGACAA*CAGAGAG ACGGTCGGGTCCACAGTTTCC* **H2:***GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTCTCGAGATGTCGCAAGACATGGCGACAT CTACCAACAGCCATGTCTTG **Target**: GACATCTACCAACAGTATCGCTC
D3	**H1**:GAGCGATA**G**GTAGGTTA**C**TTGAGCCCATGTGTATCGCTCAA**G**TAACCTAC**C**CGAGACAA*CAGAGA GACGGTCGGGTCCACAGT* **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTCTCGACTTGAGCGATACACATGGGCTCA AGTAACCTAC**C**CCATGTGTATC **Target**: GCTCAAGTAACCTACCTATCGCTC
D4	**H1**:GAGCGATA**C**GTAGGTTA**G**TTGAGCCCATGTGTATCGCTCAA**C**TAACCTAC**G**CGAGACAA*CAGAGA GACGGTCGGGTCCACAG*T **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTCTCGA**G**TTGAGCGATACACATGGGCTCA A**C**TAACCTAC**G**CCATGTGTATC **Target**: GCTCAACTAACCTACGTATCGCTC
D5	**H1**:GAGCGATACTGTTGGTAGATGTCGCCATGTCTTGCGACATCTACCAACAGCGAAACAA*CAGAGAG ACGGTCGGGTCCACAGTTTCC* **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTTTCGAGATGTCGCAAGACATGGCGACAT CTACCAACAGCCATGTCTTG **Target**: GACATCTACCAACAGTATCGCTC
D6	**H1**:GAGCGATACTGTTGGTAGATGTCGCCATGTCTTGCGACATCTACCAACAGCGAAACAA*CAGAGAG ACGGTCGGGTCCACAGTT* **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTCTCTG*TTGTTTCGAGATGTCGCAAGACATGGCGACAT CTACCAACAGCCATGTCTTG **Target**: GACATCTACCAACAGTATCGCTC
D7	**H1**:GAGCAACACTCGTGTTAGATGTCGCCATGTGTATCCGACATCTAACAC*GAGACGGTCGGGTCCAC AGTTT* **H2**:*GGAAACTGTGTCGAGTAGAGTGTGGGCTC*AGATGTCGGATACACATGGCGACATCTAACACGAG CCATGTGTATC **Target**: GACATCTAACACGAGTGTTGCTC

Table S2: RNA properties of each hairpin design.

Design	H1	H2
D1	Stem: 17 bases/ GC= 42% Loop: 8 bases/ GC= 38%	Stem: 11 bases/ GC= 54% Loop: 14 bases/ GC= 29%
D2	Stem: 17 bases/ GC= 53% Loop: 8 bases/ GC= 38%	Stem: 11 bases/ GC= 54% Loop: 14 bases/ GC= 43%
D3	Stem: 16 bases/ GC= 50% Loop: 11 bases/ GC= 46%	Stem: 12 bases/ GC= 50% Loop: 14 bases/ GC= 39%
D4	Stem: 16 bases/ GC= 50% Loop: 11 bases/ GC= 46%	Stem: 11 bases/ GC= 46% Loop: 16 bases/ GC= 50%
D5	Stem: 17 bases/ GC= 53% Loop: 8 bases/ GC= 38%	Stem: 11 bases/ GC= 54% Loop: 14 bases/ GC= 43% No 5' free region
D6	Stem: 17 bases/ GC= 53% Loop: 8 bases/ GC= 38% 4 bases shorten in Broccoli distal end	Stem: 11 bases/ GC= 54% Loop: 14 bases/ GC= 43% No 5' free region
D7	Stem: 16 bases/ GC= 50% Loop: 11 bases/ GC= 46% 4 bases shorten in Broccoli distal end Broccoli sequence as part of the stem	Stem: 11 bases/ GC= 46% Loop: 16 bases/ GC= 50%

Table S3: Sequences of the molecular beacons and their targets used in this study.

MBSgrS	GACAUCUACCAACAGUAUCGCUCACACCAAUACUCAGUCACACAGAGCGA
Target_SgrS	GGAUGUGUGACUGAGUAUUGGUGU
MBmiR21	GACAUCUACCAACAGUAUCGCUCUCAACAUCAGUCUGAUAAGCUAAGCGAUA
Target_miR21	GGAUAGCUUAUCAGACUGAUGUUGA

Table S4: Sequences of the theophylline-HHR construct. The inhibitor region is underlined.

Inhibitor 8	GGUAGAUGUCGUACAUCCAGCUGAUGAGUCCCAAAUAGGACGAAACCCAUACCAGCCGAA AGGCCCUUGGCAGGGCGUCCUGGAUUCCAGACAUCUACCAACAGUAUCGCUC
Inhibitor 9	GGGUAGAUGUCGUACAUCCAGCUGAUGAGUCCCAAAUAGGACGAAACCCAUACCAGCCGA AAGGCCCUUGGCAGGGCGUCCUGGAUUCCAGACAUCUACCAACAGUAUCGCUC
Inhibitor 10	GGGGUAGAUGUCGUACAUCCAGCUGAUGAGUCCCAAAUAGGACGAAACCCAUACCAGCCG AAAGGCCCUUGGCAGGGCGUCCUGGAUUCCAGACAUCUACCAACAGUAUCGCUC

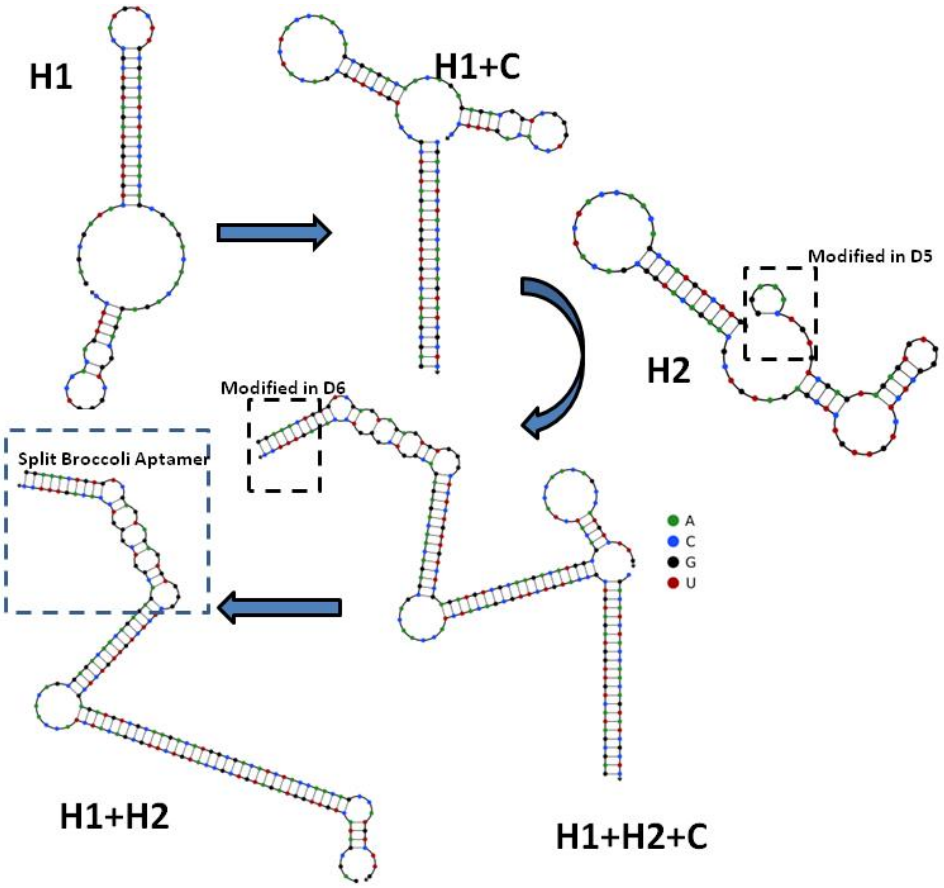

Figure S1: Predicted secondary RNA structures for Design 2 H1, H2, H1+C, H1+H2+C, and H1+H2 with NUPACK. The modified split Broccoli regions in Design 5 and 6 are shown in the black dashed boxes.

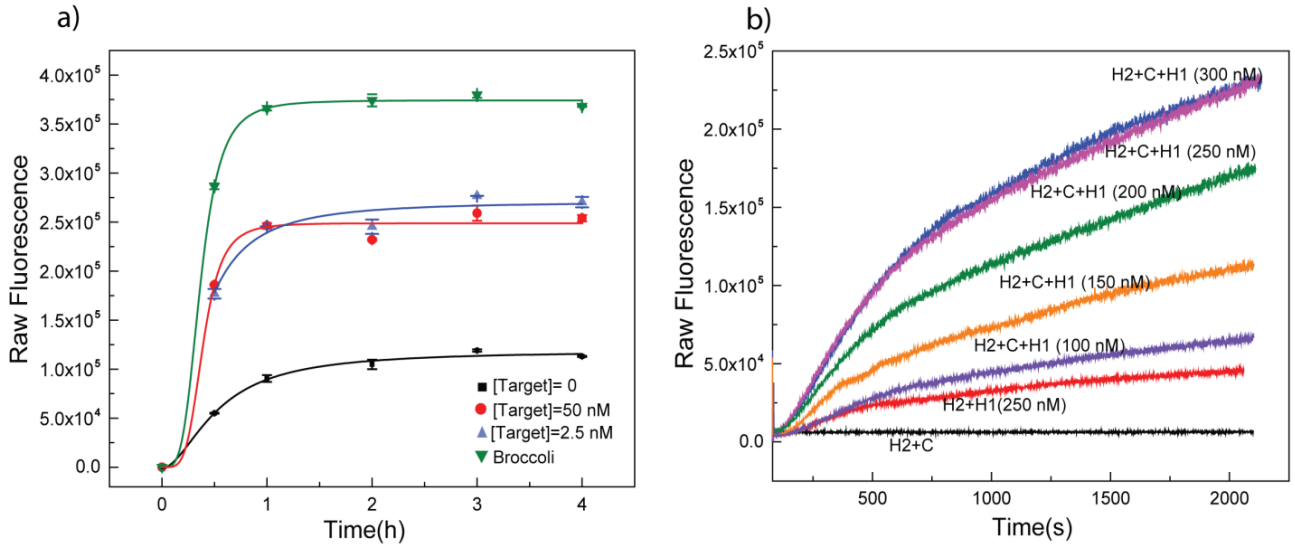

Figure S2: Determination of the assembly kinetics in D2. (a) D2 fluorescence activation as measured in a solution containing 250 nM H1 and H2 (or Broccoli), 0, 2.5 nM or 50 nM target and 5 µM DFHBI-1T. Fluorescence values of three independent replicates was shown. (b) Kinetics of the CHARGE circuit with the addition of various amounts of H1 while keeping [H2]= 250 nM, [target]= 2.5 nM, [DFHBI-1T]= 5 µM. All RNA molecules in these experiments have been PAGE gel-purified beforehand.

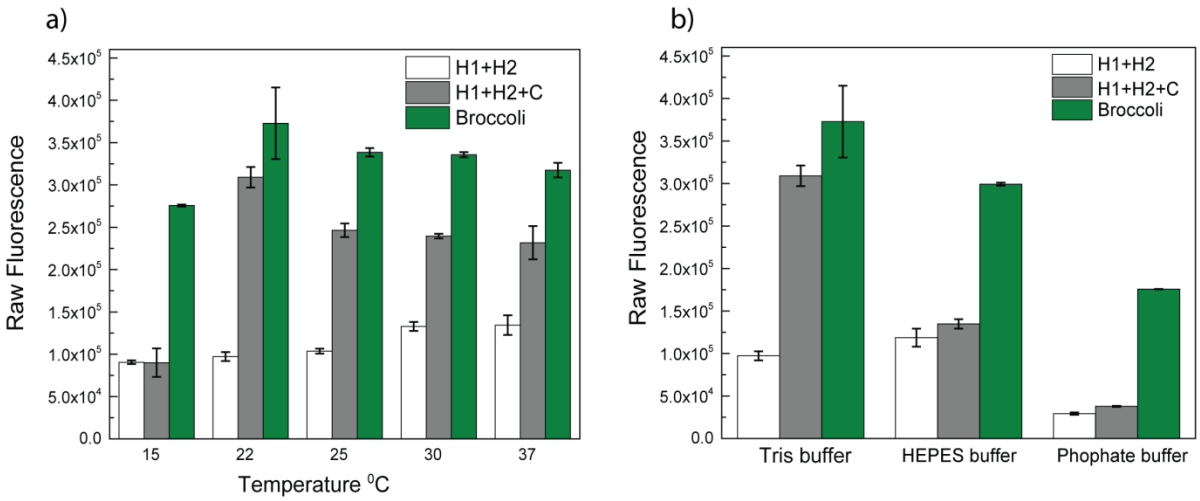

Figure S3: *In vitro* optimization of the D2 CHARGE system. (a) Temperature effect on the CHA assembly efficiency. An optimal signal-to-background ratio (3.0-fold) was observed at 22°C. Spectra was measured in solution consisting of 10 mM Tris, 5 mM $MgCl_2$, 100 mM KCl, and 10 mM NaCl at pH= 7.5. (b) The effect of different buffer conditions on D2 CHA assembly efficiency at 22°C. For all these measurements, 250 nM H1 and H2 (or Broccoli), 2.5 nM target and 5 µM DFHBI-1T was used and all these RNA molecules were PAGE gel-purified beforehand.

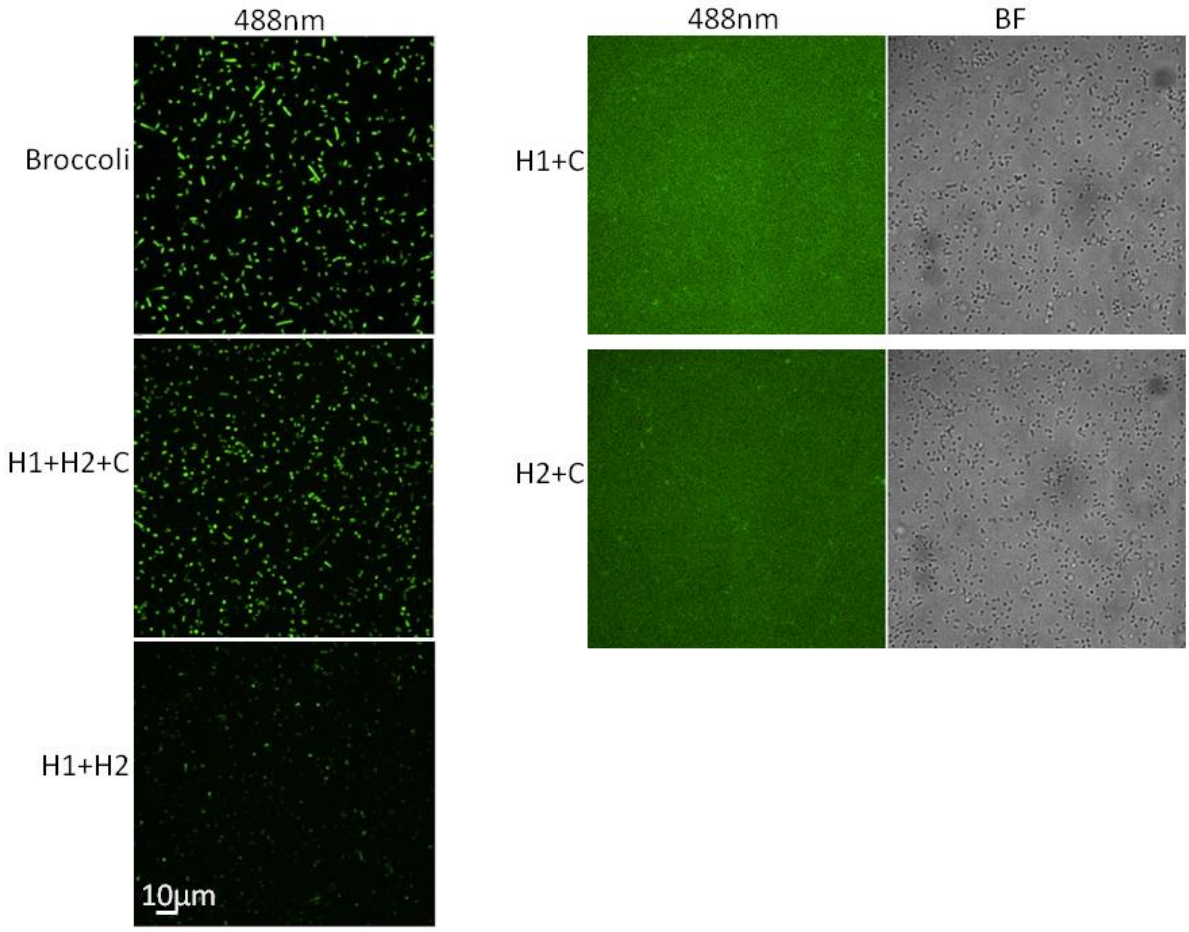

Figure S4: Confocal fluorescence imaging in live BL21 (DE3)* cells. H1 and H2 hairpins were expressed from a pETDuet vector, and target (C) was cloned in a pCDFDuet vector. After 2 h IPTG induction, 200 μM DFHBI-1T was added 30 min before imaging. The fluorescence of catalyst with H1 or H2 hairpin alone being coexpressed was too low to be determined.

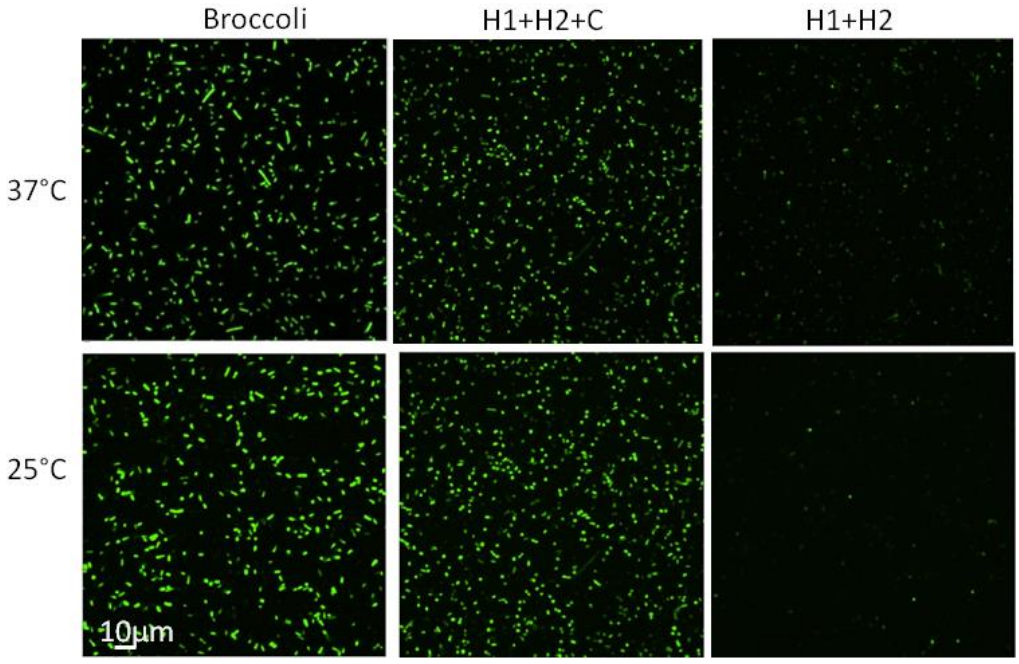

Figure S5: Temperature effect on the cellular fluorescence of CHARGE. Confocal fluorescence imaging in live BL21 (DE3)* cells expressing Broccoli aptamer, H1+H2 and H1+H2+C at 25°C and 37°C. After 2 h IPTG induction, 200 μM DFHBI-1T was added 30 min before imaging. It appears that the cellular fluorescence of Broccoli and CHARGE can be observed at both temperature. Slightly enhanced cellular fluorescence as well as improved signal-to-background ratio of CHARGE was observed at 25°C compared to that of 37°C, which is consistent with our *in vitro* data as shown in Figure S3.

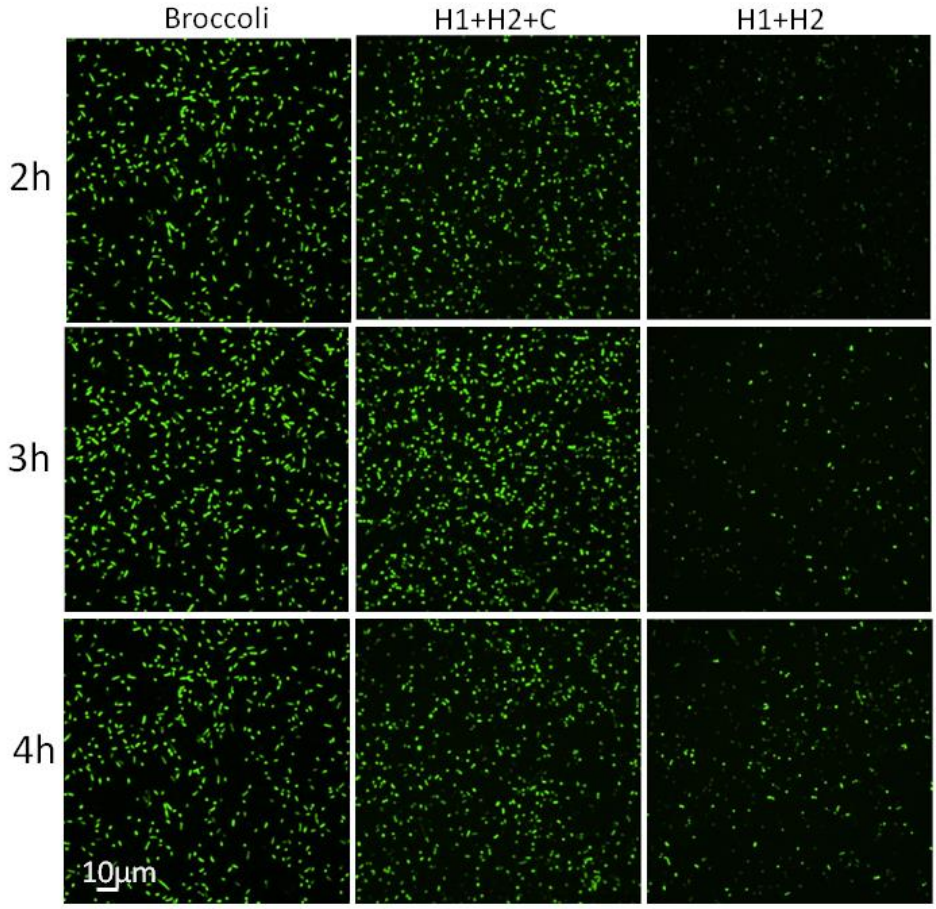

Figure S6: The cellular stability and leakage of CHARGE after a long incubation period. Confocal fluorescence imaging in live BL21 (DE3)* cells expressing Broccoli aptamer, H1+H2, or H1+H2+C at different time point (2 h, 3 h, and 4 h) after 2 h IPTG induction. 200 μM DFHBI-1T was added 30 min before imaging. It appears that the background signal leakage from H1+H2 was obviously increased at longer incubation time. Relatively stable CHARGE signal can be still visualized after a total of 6 h incubation since the addition of IPTG.

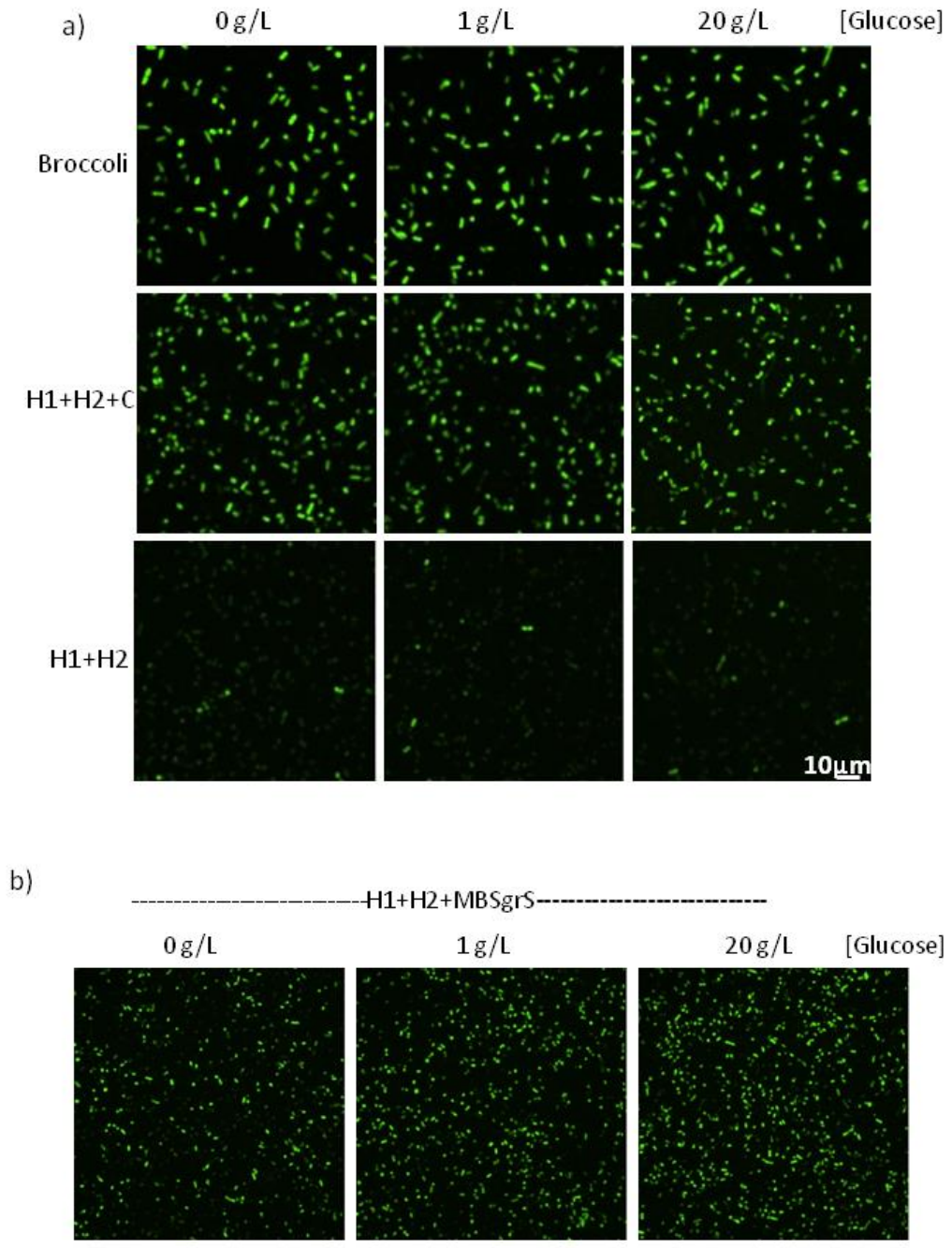

Figure S7: (a) Confocal fluorescence imaging in live BL21 (DE3)* cells expressing Broccoli, H1+H2+C, and H1+H2 after adding different concentrations (0, 1 g/L, 20 g/L) of glucose. After 2 h IPTG induction, 200 µM DFHBI-1T was added 30 min before imaging, and the images were taken 45 min after adding DFHBI-1T. (b) Images of cells expressing H1+H2+MBSgrS (SgrS-targeting molecular beacon) 2.5 h after adding different concentrations of glucose.

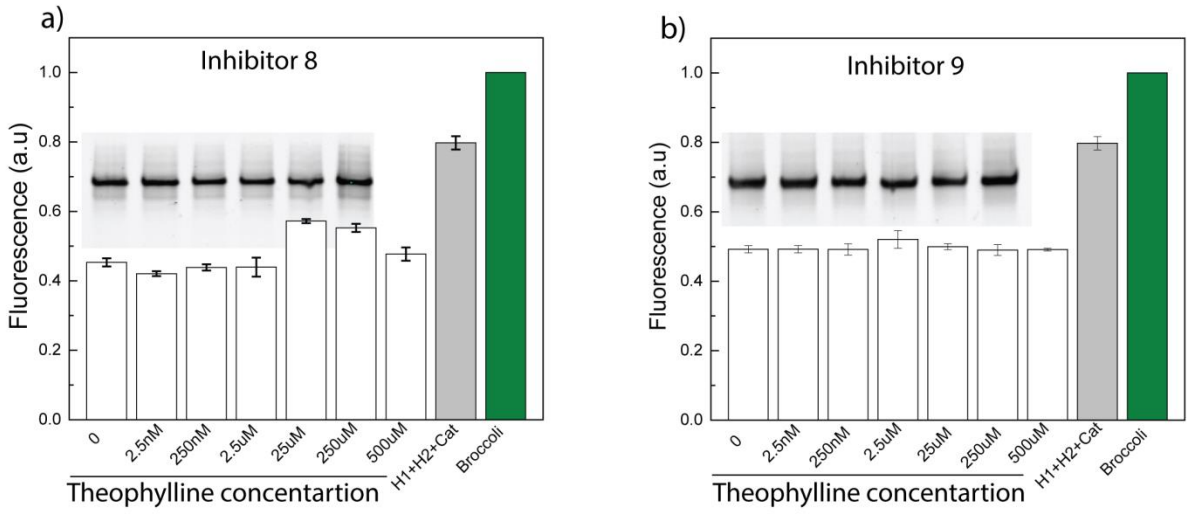

Figure S8: Fluorescence assay and 20% denaturing PAGE gel data for theophylline-HHR construct with (a) Inhibitor 8 and (b) Inhibitor 9, respectively. Inhibitor 8 exhibits some fluorescence enhancement while Inhibitor 9 was almost completely inactive.

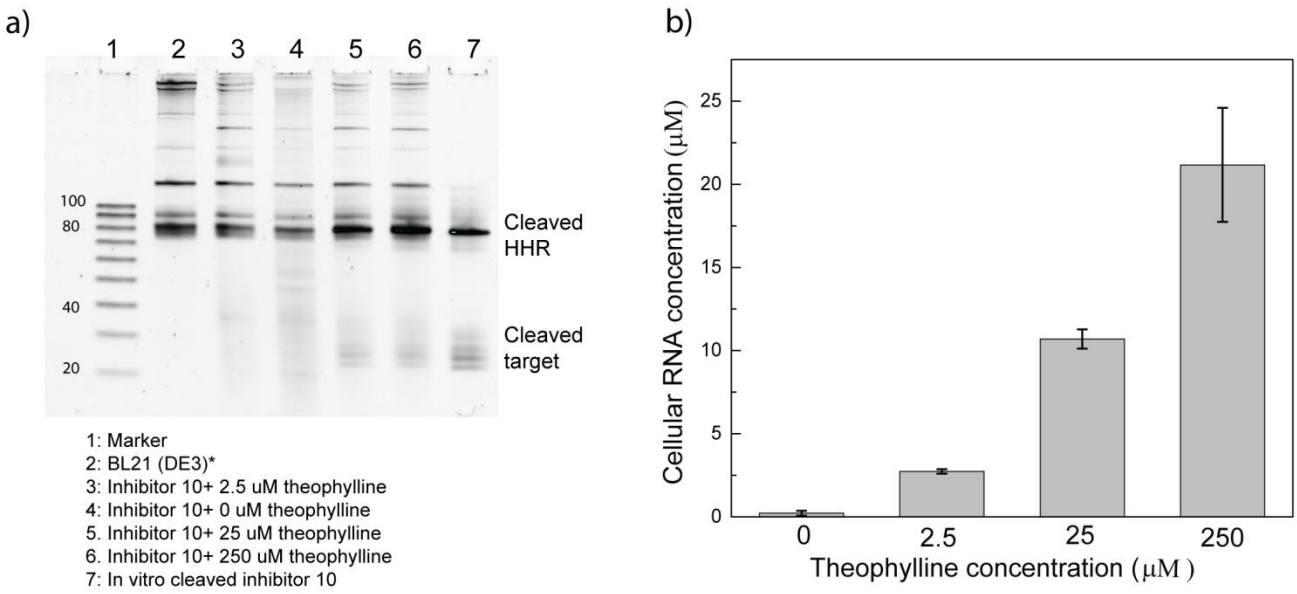

Figure S9: Theophylline-regulated cellular RNA levels in a hammerhead ribozyme system. (a) Gel analysis of cellular RNAs that isolated from BL21 (DE3)* cells expressing a pETDuet-theophylline-HHR plasmid. After induction with IPTG for 1 h, cells were treated with different concentrations of theophylline, 0, 2.5 µM, 25 µM, and 250 µM, for 30 min (Lane 3 – 6). Cellular RNAs were isolated using TRIzol Reagent (Invitrogen) and resolved in a 10% denaturing PAGE gel. (b) Theophylline-dependent activation of the cellular concentrations of target RNA. The total number of cells being isolated was calculated based on the OD_{600} value before the RNA extraction. Total concentration of the isolated RNAs was determined using a NanoDrop One UV-Vis spectrophotometer. Mass percentage of the cleaved HHR fraction in isolated total RNAs was calculated based on the integrated SYBR Gold staining intensity analysis using the Image J software. Background signal from the integrated staining intensity at the same band positions in Lane 2, i.e., BL21 (DE3)* cells without theophylline-HHR plasmid, was subtracted in this analysis. Error bars indicated the S.E.M. values from the analysis of three separated gels. Under our experimental condition, on average a 0.11 – 17 µM concentration of cellular target RNA was generated in the presence of 0 – 250 µM theophylline.

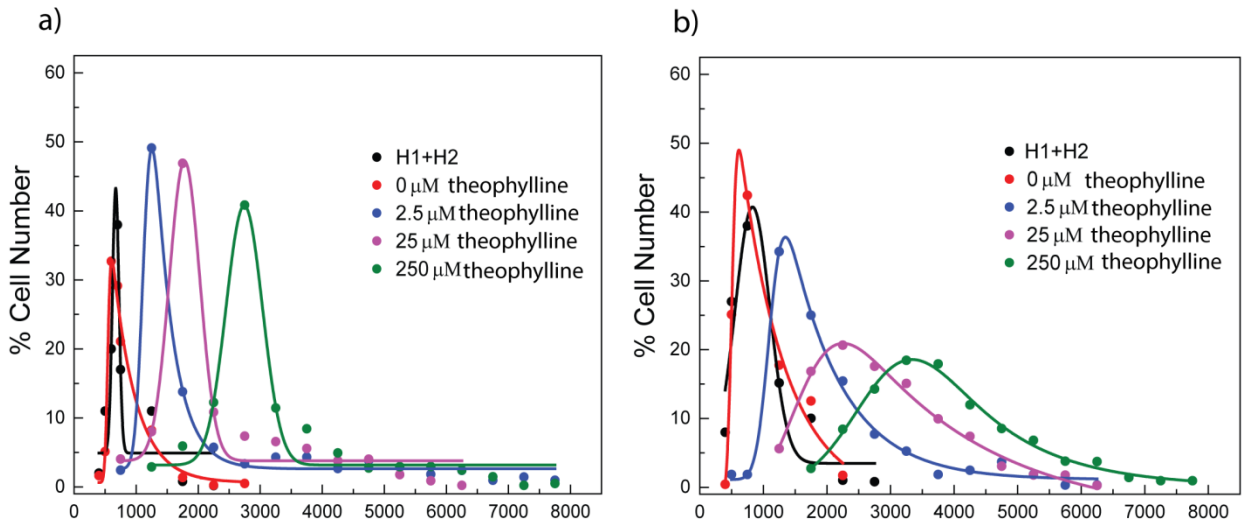

Figure S10: Distributions of cellular fluorescence level after the addition of different concentrations of theophylline. BL21 (DE3)* cells were induced by IPTG for 1 h before adding 0 – 250 μM concentrations of theophylline for an additional (a) 30 min or (b) 1 h before imaging. Here, for each experimental condition, individual cells were binned according to their brightness. The percentage of cells in each bin was plotted. A total of 200 cells were measured in each case from three experimental replicates. Each plot was then fitted based on a Gaussian function. The mean value of cellular fluorescence and standard deviation was calculated based on the Gaussian distribution. A calibration curve was plotted based on the cellular fluorescence vs. target RNA concentration as shown in Fig. S9. Based on the calibration curve, the limit of detection under this circumstance was measured to be around 110 copies of target RNA per cell.

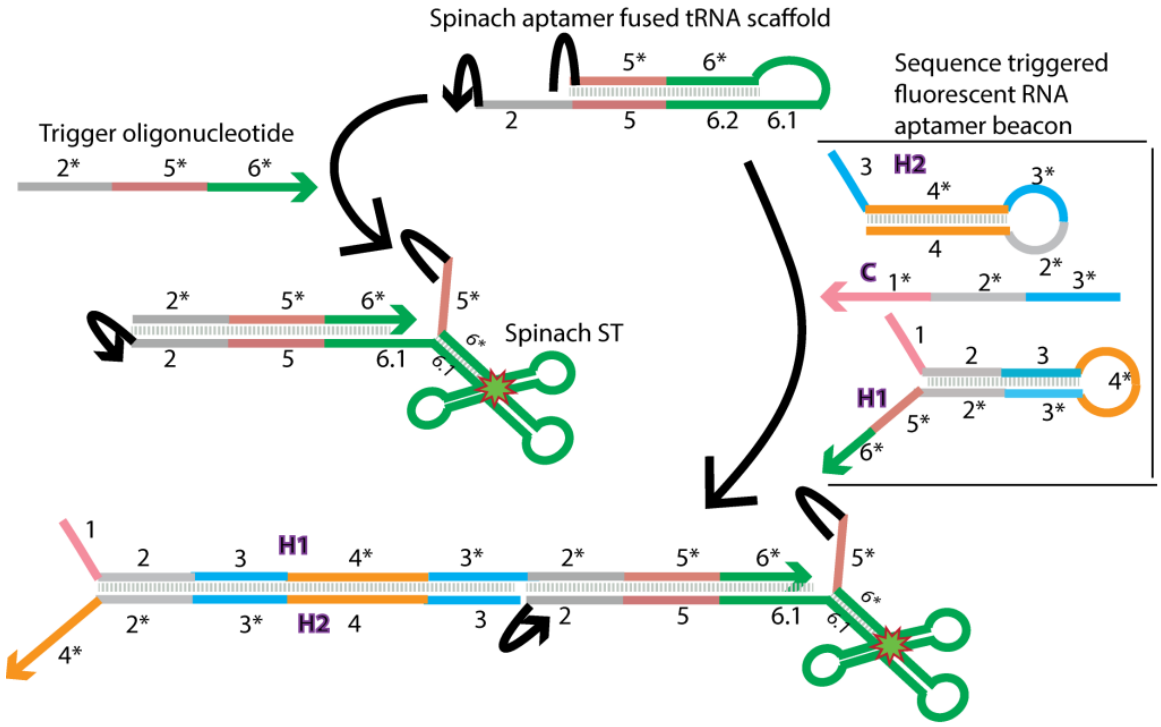

Figure S11: Schematic of Spinach aptamer-based CHA system designed by the Ellington lab [2] . Spinach is designed in such a way that it is embedded within a tRNA scaffold (black region). Trigger oligonucleotide will hybridize with tRNA-Spinach through a toehold-mediated strand displacement mechanism. In another event, H1 and H2 hybridize by CHA assembly and resulting product will hybridize with tRNA-Spinach scaffold. Finally, the Spinach fluorescence was recovered.

References

[1] R. L. Strack, W. Song, and S. R. Jaffrey, "Using Spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria," *Nat. Protoc.*, vol. 9, no. 1, pp. 146–155, 2014.

[2] S. Bhadra and A. D. Ellington, "Design and application of cotranscriptional non-enzymatic RNA circuits and signal transducers," *Nucleic Acids Res.*, vol. 42, no. 7, pp. 1–16, 2014.