US Patent & Trademark Office
Patent Public Search | Text View

United States Patent	12391948
Kind Code	B2
Date of Patent	August 19, 2025
Inventor(s)	Jaffrey; Samie R. et al.

RNA-regulated fusion proteins and methods of their use

Abstract

The present disclosure is directed to RNA-regulated fusion proteins comprising a protein of interest and an RNA-regulated destabilization domain. Also disclosed are RNA aptamers that bind specifically to a RNA-regulated destabilization domain. Nucleic acid molecules encoding the RNA-regulated fusion proteins and RNA aptamers and methods of use thereof are also disclosed.

Inventors: **Jaffrey; Samie R. (New York, NY), Wu; Jiahui (New York, NY)**

Applicant: **CORNELL UNIVERSITY** (Ithaca, NY)

Family ID: **1000008767720**

Assignee: **CORNELL UNIVERSITY (Ithaca, NY)**

Appl. No.: **17/637940**

Filed (or PCT Filed): **August 31, 2020**

PCT No.: **PCT/US2020/048781**

PCT Pub. No.: **WO2021/042050**

PCT Pub. Date: March 04, 2021

Prior Publication Data

Document Identifier	Publication Date
US 20220290161 A1	Sep. 15, 2022

Related U.S. Application Data

us-provisional-application US 62894651 20190830

Publication Classification

Int. Cl.: **C12N15/62** (20060101); **A61K31/7105** (20060101); **C07K19/00** (20060101); **C12N5/071** (20100101); **C12N9/02** (20060101); **C12N15/115** (20100101); **C12N15/52** (20060101); **C12N15/85** (20060101); **C12Q1/6816** (20180101)

U.S. Cl.:

CPC **C12N15/62** (20130101); **C12N5/0602** (20130101); **C12N15/115** (20130101); **C12N15/52** (20130101); **C12N15/85** (20130101); **C12Q1/6816** (20130101); **C12Y111/01011** (20130101); **C12Y113/12013** (20130101); **C12Y201/01043** (20130101); **C12Y603/0401** (20130101); C07K2319/60 (20130101); C07K2319/61 (20130101); C07K2319/85 (20130101); C12N2310/16 (20130101); C12N2800/107 (20130101)

Field of Classification Search

USPC: None

References Cited

U.S. PATENT DOCUMENTS

Patent No.	Issued Date	Patentee Name	U.S. Cl.	CPC
9388418	12/2015	Rossi et al.	N/A	N/A
2008/0300187	12/2007	Sauvageau	435/375	C07K 14/47
2012/0213737	12/2011	Zhu et al.	N/A	N/A
2014/0255361	12/2013	Wandless et al.	N/A	N/A
2019/0185934	12/2018	Jaffrey et al.	N/A	N/A

OTHER PUBLICATIONS

International Search Report and Written Opinion for corresponding International Application No. PCT/US2020/048781 (mailed Feb. 8, 2021). cited by applicant

Wu et al., "Live Imaging of mRNA using RNA-Stabilized Fluorogenic Proteins," Nat. Method. 16(9):862-865 (2019). cited by applicant

Song et al., "Imaging RNA Polymerase III Transcription Using a Photostable RNA-Fluorophore Complex," Nat. Chem. Biol. 13(11):1187-1194 (2017). cited by applicant

Extended European Search Report for European Application No. 20950485.1 (mailed Mar. 4, 2025). cited by applicant

Tok et al., "A Comparative Binding Study of Modified Bovine Immunodeficiency Virus TAR RNA Against its TAT Peptide," Bioorg. Med. Chem. Lett. 14(24):6101-6105 (2004). cited by applicant

Bonger et al., "Small-Molecule Displacement of a Cryptic Degron Causes Conditional Protein Degradation," Nat. Chem. Biol. 7(8):531-537 (2011). cited by applicant

Battiste et al., "α Helix-RNA Major Groove Recognition in an HIV-1 Rev Peptide-RRE RNA Complex," Science 273:1547-1551 (1996). cited by applicant

Legault et al., "NMR Structure of the Bacteriophage λ N Peptide/boxB RNA Complex: Recognition of a GNRA Fold by an Arginine-Rich Motif," Cell 93(2):289-299 (1998). cited by applicant

Yamamoto et al., "A Novel RNA Motif that Binds Efficiently and Specifically to the Tate Protein of HIV and Inhibits the Trans-Activation by Tate of Transcription in vitro and in vivo," Genes to Cells 5:371-388 (2000). cited by applicant

Rakhit et al., “Evaluation of FKBP and DHFR based Destabilizing Domains in *Saccharomyces cerevisiae*,” Bioorg. Med. Chem. Lett. 21(17):4965-4968 (2011) [Author manuscript]. cited by applicant
Iwamoto et al., “A General Chemical Method to Regulate Protein Stability in the Mammalian Central Nervous System,” Chem. Biol. 17(9):981-988 (2010) [Author Manuscript]. cited by applicant
Athanassiou et al., “Structural Mimicry of Retroviral Tat Proteins by Constrained β-Hairpin Peptidomimetics: Ligands with High Affinity and Selectivity for Viral TAR RNA Regulatory Elements,” J. Am. Chem. Soc. 126(22):6906-6913 (2004). cited by applicant
Chen et al., “A Peptide Interaction in the Major Groove of RNA Resembles Protein Interactions in the Minor Groove of DNA,” Proc. Natl. Acad. Sci. USA 92:5077-5081 (1995). cited by applicant
Chen and Frankel, “An RNA-Binding Peptide from Bovine Immunodeficiency Virus Tat Protein Recognizes an Unusual RNA Structure,” Biochemistry 33:2708-2715 (1994). cited by applicant
Koren et al., “The Eukaryotic Proteome is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons,” Cell 173:1622-1635 (2018). cited by applicant
Smith et al., “Altering the Context of an RNA Bulge Switches the Binding Specificities of Two Viral Tat Proteins,” Biochemistry 37:10808-10814 (1998). cited by applicant

---

Primary Examiner: Rao; Manjunath N

Assistant Examiner: Lee; Jae W

Attorney, Agent or Firm: Troutman Pepper Locke LLP (Rochester)

Background/Summary

(1) This application is a national stage application under 35 U.S.C. § 371 of International Application No. PCT/US2020/048781, filed Aug. 31, 2020, which claims priority benefit of U.S. Provisional Patent Application Ser. No. 62/894,651 filed Aug. 30, 2019, which is hereby incorporated by reference in its entirety.

FIELD OF THE DISCLOSURE

(1) This present disclosure relates to RNA-regulated fusion proteins and methods of their use.

BACKGROUND

(2) Fluorogenic RNA aptamers are RNA aptamers that bind otherwise nonfluorescent molecules and switch them to a fluorescent form. These fluorogenic dyes can be applied to cells, enabling RNAs tagged with these fluorogenic aptamers to be imaged using fluorescence microscopy (Paige et al., “RNA Mimics of Green Fluorescent Protein,” *Science* 333:642-646 (2011) and Braselmann et al., “A Multicolor Riboswitch-Based Platform for Imaging of RNA in Live Mammalian Cells,” *Nat. Chem. Biol.* 14:964-971 (2018)). However, few fluorogenic aptamers have been developed since there are not many fluorogenic dyes that meet the criteria required for use in live cells. For example, most dyes show nonspecific fluorescence activation by cellular lipids or DNA (custom character G., “The Fluorescence of Dye—Nucleic Acid Complexes,” *Journal of Luminescence* 22:221-265 (1981) and Fam et al., “Recent Advances in Fluorescent Probes for Lipid Droplets,” *Materials* (*Basel*) 11 (2018)). This nonspecific binding leads to background fluorescence that obscures the fluorescence of the RNA-dye complexes. Another problem is that the fluorogenic dyes are not genetically encoded and therefore need to be added exogenously for RNA imaging. A genetically encoded conditionally fluorescent dye would provide a simple alternative to the use of fluorogenic RNA aptamers.

(3) The present disclosure is directed to overcoming deficiencies in the art.

SUMMARY

(4) A first aspect of the disclosure relates to a nucleic acid molecule encoding an RNA-regulated fusion protein. The nucleic acid molecule includes: a first nucleic acid sequence encoding a protein of interest and a second nucleic acid sequence encoding an RNA-regulated destabilization domain, where the second nucleic acid sequence is operably coupled to the first nucleic acid sequence.

(5) Another aspect of the disclosure relates to a nucleic acid molecule encoding a lentiviral transactivator of transcription (Tar) RNA aptamer sequence.

(6) A further aspect of the disclosure relates to an RNA-regulated fusion protein comprising a protein of interest and an RNA-regulated destabilization domain.

(7) Yet another aspect of the disclosure relates to a molecular complex comprising: an RNA-regulated fusion protein comprising (i) a protein of interest and (ii) an RNA-regulated destabilization domain; and an RNA aptamer bound specifically to the RNA-regulated destabilization domain.

(8) Another aspect of the invention relates to a method of imaging RNA in a cell. This method involves providing a first vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a fluorescent protein, a bioluminescent protein, or an enzyme fused to an RNA-regulated destabilization domain; providing a second vector encoding an RNA molecule comprising (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; transfecting a host cell with the first vector and the second vector; and imaging said transfected cells.

(9) Yet another aspect of the invention relates to a method of imaging RNA in a cell. This method involves providing a vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a fluorescent protein, a bioluminescent protein, or an enzyme fused to an RNA-regulated destabilization domain; transfecting a host cell with the first vector; contacting said transfected cell with an RNA molecule comprising (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; and imaging said contacted cells.

(10) A further aspect of the invention relates to a method of selectively modifying an RNA-binding protein. This method involves providing a first expression vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises an enzyme fused to an RNA-regulated destabilization domain; providing a second expression vector encoding (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; transfecting a host cell with the first and second expression vectors; and allowing the enzyme to be expressed, where the expressed enzyme selectively modifies a protein that binds to the RNA sequence of interest.

(11) Another aspect of the invention relates to a method of regulating expression of an RNA-stabilized protein of interest. This method involves providing a first vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a protein of interest fused to an RNA-regulated destabilization domain; providing a second vector encoding an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; providing a host cell comprising a functional ubiquitination system; transfecting the host cell with the first and second expression vectors; and expressing the first and second expression vectors within the host cell, where said expressing the first and second expression vectors regulates proteomic stability of the RNA-regulated fusion protein; and where, in the absence of any expressed RNA aptamer sequence in the host cell, the RNA-regulated destabilization domain promotes degradation of the RNA-regulated fusion protein by the ubiquitination system; and where the RNA-regulated fusion protein is stabilized by the expressed RNA aptamer sequence.

(12) Another aspect of the invention relates to a method of regulating expression of an RNA-stabilized protein of interest. This method involves providing a first vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a protein of interest fused to an RNA-regulated destabilization domain; providing a second vector encoding an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; providing a mammalian cell lysate or solution comprising (i) a ubiquitin ligase, (ii) proteosomal degradation machinery, (iii) transcriptional machinery, and (iv) translational machinery; contacting the mammalian cell lysate or solution with the first and second expression vectors; and expressing the first and second expression vectors, where said expressing the first and second expression vectors regulates proteomic stability of the RNA-regulated fusion protein; and where, in the absence of any expressed RNA aptamer sequence in the cell lysate or solution, the RNA-regulated destabilization domain promotes degradation of the RNA-regulated fusion protein by the proteosomal degradation system; and where the RNA-regulated fusion protein is stabilized by the expressed RNA aptamer sequence.
(13) Another aspect of the present application relates to a treatment method. This method involves contacting a cell with an RNA aptamer, where upon said contacting, the aptamer interacts with an RNA-regulated destabilization domain fused to a protein of interest in the cell to stabilize the protein of interest in the cell.
(14) Another aspect of the present invention relates to a treatment method. This method involves contacting a cell with a vector according to the present application under conditions effective to express an RNA molecule as described herein to treat the cell.
(15) The examples described herein below demonstrate the use of RNA-regulated fluorescent fusion proteins whose fluorescence is stabilized by RNA aptamers. In some embodiments, the RNA-regulated fluorescent fusion proteins are highly unstable until they bind RNA aptamers inserted in mRNAs, resulting in fluorescent RNA-protein complexes that enable live imaging of mRNA in living cells. In some embodiments, the technology described herein is an imaging system that bypasses the limitations of using fluorogenic RNA aptamers and conditionally fluorescent small molecule dyes for imaging. In some embodiments, this is achieved by engineering a peptide degron sequence whose activity can be regulated by an RNA aptamer. When fused to a fluorescent protein, this peptide degron sequence can send the fluorescent protein to degradation. However, this degradation function of the peptide degron is impeded when bound to a specific RNA aptamer sequence. In some embodiments, a peptide degron sequence causes rapid degradation of the unbound fluorescent proteins when expressed in mammalian cells. This is different from previous methods. In some embodiments, methods described herein utilize an RNA aptamer sequence that can effectively abrogate the degradation function of the peptide degron once they are bound. This is also different from previous methods. Methods described herein enable fluorescent proteins and other proteins to carry out their native function only when they are bound to a specific RNA sequence. In the case of enhanced yellow fluorescent protein (EYFP), a 38 fold fluorescent enhancement was observed when bound to the engineered RNA aptamer described herein.

Description

BRIEF DESCRIPTION OF THE DRAWINGS
(1) FIGS. **1**A-**1**C show the design and optimization of an RNA-regulated protein destabilization domain. FIG. **1**A is a schematic drawing of a Pepper RNA-regulated protein destabilization domain, tDeg. tDeg is a bifunctional peptide that includes the Tat peptide, which is capable of binding to the Pepper RNA aptamer, and the previously described C-terminal Arg-Arg-Arg-Gly degron (Bonger et al., “Small-Molecule Displacement of a Cryptic Degron Causes Conditional Protein Degradation,” *Nat. Chem. Biol.* 7:531-7 (2011), which is hereby incorporated by reference

in its entirety). When fused to a protein of interest, tDeg causes protein degradation. However, the protein destabilization function of tDeg is impeded when it binds to the Pepper RNA aptamer. Amino acids Arg-Gly, highlighted in a black box, are appended to the C-terminus of Tat to make the full Arg-Arg-Arg-Gly degron. FIG. **1**B demonstrates that Pepper RNA stabilizes EYFP fused to tDeg in cells. To test whether tDeg functions as an RNA-regulated destabilization domain, EYFP-tDeg was coexpressed with different circular RNAs, and the yellow fluorescence in HEK293T cells was imaged. Without circular wild-type TAR RNA or its variants, cells coexpressing EYFP-tDeg and the circular control RNA only showed minimal fluorescence above background fluorescence. Cells exhibit yellow fluorescence only when circular wild-type TAR RNA, TAR Variant-1, or TAR Vairnat-2 (named Pepper) was coexpressed. Notably, higher yellow fluorescence signals were observed in the cytosol compared to the nucleus when EYFP-tDeg was coexpressed with the circular wild-type TAR RNA or its variants. This is consistent with the cytosolic expression of small circular RNAs using the Tornado expression system (Litke & Jaffrey, "Highly Efficient Expression of Circular RNA Aptamers in Cells Using Autocatalytic Transcripts," *Nat. Biotechnol.* 37:667-675 (2019), which is hereby incorporated by reference in its entirety). All cells were stained with Hoechst dye. Scale bar, 40 μm. FIG. **1**C shows the summary data of normalized fluorescence of untransfected HEK293T cells, or HEK293T cells expressing EYFP or EYFP-tDeg with different RNAs as in (FIG. **1**B). Total cellular yellow fluorescence of individual cells is plotted (n=4 independent cell cultures). Values are means±s.d. ****P.sub.circular wild-type TAR=7.9×10.sup.−113; ****P.sub.circular TAR Variant-1=2.1×10.sup.−117; ****P.sub.circular TAR Variant-2=1.7×10.sup.−115 by one-way ANOVA.

(2) FIGS. **2**A-**2**B are schematic illustrations showing the design of tDeg, an RNA-regulated destabilization domain. Shown is a structural representation of how TAR binds to the tDeg, and may therefore obstruct recognition of the Arg-Arg-Arg-Gly degradation-inducing signal. RNA is depicted in grey, and peptide sequence is shown letters of the polypeptide chain. A schematic representation of RNA binding to the tDeg sequence is shown in FIG. **2**A. Here, a bifunctional peptide sequence, called tDeg, that functions both as a destabilization domain and as a binding site for the bovine immunodeficiency virus TAR RNA (in grey) was designed. Knowing that the TAR RNA binds to specific amino acids in the Tat peptide including the two C-terminal arginines, an Arg-Gly (highlighted in a black box) was added to the C-terminus of the Tat peptide to make the full Arg-Arg-Arg-Gly degron. When the TAR RNA binds to this bifunctional domain, it impedes the function of the destabilization domain by sterically blocking recognition of the Arg-Arg-Arg-Gly degron by proteasomal machinery. The structure model (FIG. **2**B) of the Tat-TAR complex shows that the first two arginines of the Arg-Arg-Arg-Gly degron would be inaccessible to any Arg-Arg-Arg-Gly-binding protein that mediates its degradation. The additional Arg-Gly residues are modeled into the C-terminus of Tat in a black box in FIG. **2**B. The structure representation in FIG. **2**B is based on the NMR structure of the bovine immunodeficiency virus Tat-TAR complex (PDB entry: 1BIV) (Puglisi et al., "Solution Structure of a Bovine Immunodeficiency Virus Tat-TAR Peptide-RNA Complex," *Science* 270:1200-3 (1995), which is hereby incorporated by reference in its entirety).

(3) FIGS. **3**A-**3**B demonstrate that tDeg confers protein instability to EYFP by proteasomal degradation. In FIG. **1**B, it was shown that tDeg confers protein instability to EYFP. However, the lack of yellow fluorescence of EYFP-tDeg in FIG. **1**B could be due to protein misfolding or aggregation. In FIG. **3**A, whether the lack of yellow fluorescence of EYFP-tDeg is due to proteasomal degradation was examined. In these experiments, HEK293T cells were transiently transfected with a plasmid expressing EYFP-tDeg. These cells were then treated with vehicle (DMSO) or a proteasome inhibitor (10 μM MG132) for 7 hours, respectively. When treated with vehicle (DMSO), minimal yellow fluorescence was detected. This result is consistent with the result from FIG. **1**B. However, when proteasome activity was inhibited by treatment of 10 μM MG132 for 7 hours, the yellow fluorescence of EYFP-tDeg was restored. Thus, this confirmed that

the tDeg tag markedly reduces the stability of EYFP by inducing its proteasomal degradation. All cells were stained with Hoechst dye. Scale bar, 40 μm. In FIG. **3**B, normalized total cellular yellow fluorescence of individual cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P=5.6×10.sup.−36 by unpaired two-tailed Student's t-test.

(4) FIGS. **4**A-**4**B demonstrate that engineered TAR variants' higher efficiency in stabilizing EYFP-tDeg proteins is not due to expression differences in EYFP-tDeg mRNA or the circular TAR RNAs. In FIGS. **1**B and **1**C, it was shown that circular wild-type TAR, Variant-1, and Variant-2 showed 24-fold, 36-fold, and 38-fold fluorescence increases, respectively. However, the improved efficiency in stabilizing EYFP-tDeg protein could be due to uneven expression levels of the EYFP-tDeg mRNA, or the uneven expression levels of the circular TAR RNA variants. Here, the relative expression of EYFP-tDeg mRNA (FIG. **4**A) and the relative expression of circular TAR RNA variants (FIG. **4**B) was compared. In these experiments, HEK293T cells were transiently transfected with a plasmid expressing EYFP-tDeg and the corresponding circular TAR RNA variant as shown in FIGS. **1**B and **1**C. Total RNA was extracted by TRIzol® extraction. EYFP-tDeg mRNA expression level was quantified using RT-qPCR. Each circular TAR RNA variant's expression level was quantified by running the extracted total RNA on a TBE-Urea gel followed by SYBR™ Gold nucleic acid gel staining. These results show that there is no significant expression difference in the EYFP-tDeg mRNA or the circular TAR RNA variants. Thus, this confirms that the engineered circular TAR RNA variants indeed show higher efficiency in stabilizing tDeg-tagged EYFP. Data were collected from two independent cell cultures. Values are means±s.d.

(5) FIGS. **5**A-**5**G demonstrate that tDeg can be regulated by the Pepper RNA aptamer in diverse mammalian cell types. In FIGS. **1**A-**1**C, it was shown that EYFP-tDeg can be regulated by the Pepper RNA aptamer in HEK293T cells. Here, whether tDeg can be regulated by the Pepper RNA aptamer in various mammalian cell types was examined (FIG. **5**A). In these experiments, U2OS cells (FIG. **5**B, FIG. **5**E), COS-7 cells (FIG. **5**C, FIG. **5**F), or HeLa cells (FIG. **5**D, FIG. **5**G) were transiently expressed EYFP-tDeg with and without the circular Pepper RNA aptamer, respectively. In each case, cells showed low or undetectable levels of yellow fluorescence without the circular Pepper RNA aptamer. The yellow fluorescence of EYFP-tDeg was only restored when the circular Pepper RNA aptamer was coexpressed. Thus, tDeg can be regulated by the Pepper RNA aptamer in diverse mammalian cell types. All cells were stained with Hoechst dye. Scale bar, 20 μm. Normalized total cellular fluorescence (FIGS. **5**E, **5**F, and **5**G) of individual cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P.sub.U2OS=5.7×10.sup.−59; ****P.sub.COS-7=1.6×10.sup.−46; ****P.sub.HeLa=2.0×10.sup.−139 by unpaired two-tailed Student's t-test.

(6) FIGS. **6**A-**6**G demonstrate that tDeg confers Pepper RNA-dependent regulation to diverse proteins. To test whether Pepper RNA stabilizes different proteins fused to tDeg, HEK293T cells expressing mNeonGreen (FIG. **6**B, FIG. **6**E), mCherry (FIG. **6**C, FIG. **6**F), and the luciferase NanoLuc (FIG. **6**D, FIG. **6**G) fused to a C-terminal tDeg tag with and without circular Pepper RNA (FIG. **6**A) were imaged, respectively. In each case, there was a considerable increase of fluorescence (FIG. **6**E, FIG. **6**F) or bioluminescence (FIG. **6**G) of the tDeg-tagged protein only when circular Pepper RNA was coexpressed in cells. For detecting bioluminescence, cells were incubated in media with furimazine (from Promega Nano-Glo® Luciferase Assay System, diluted 100×) and imaged using a 460±25 nm emission filter cube. All cells were stained with Hoechst dye. Scale bar, 40 μm. Normalized total cellular fluorescence (FIG. **6**E and FIG. **6**F) or bioluminescence (FIG. **6**G) of individual cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P.sub.mNeonGreen-tDeg=1.1×10.sup.−123; ****P.sub.mCherry-tDeg=3.0×10.sup.−131; ****P.sub.NanoLuc-tDeg=1.7×10.sup.−120 by unpaired two-tailed Student's t-test.

(7) FIGS. **7**A-**7**G demonstrate that tDeg confers Pepper RNA-dependent regulation to diverse proteins. In FIGS. **6**A-**6**G, it was shown that tDeg confers Pepper RNA-dependent regulation of

different fluorescent proteins and the luciferase, NanoLuc (Hall et al., "Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate," *ACS Chem. Biol.* 7:1848-57 (2012), which is hereby incorporated in its entirety). Whether tDeg confers Pepper-dependent regulation to proteins with different functions and localizations in cells was tested here (FIG. **7**A). In these experiments, HEK293T cells transiently expressed EGFP-TetR-tDeg (FIG. **7**B, FIG. **7**E), EGFP-EZH2-tDeg (FIG. **7**C, FIG. **7**F), or mCherry-NF-κB-tDeg (FIG. **7**D, FIG. **7**G), with and without the circular Pepper RNA aptamer, respectively. In each case, proteins were nearly undetectable unless coexpressed with the circular Pepper RNA. Furthermore, protein localization of these proteins without tDeg and the circular Pepper RNA was compared to their stabilized counterparts by tDeg and circular Pepper RNA. It was observed that EGFP-TetR-tDeg with circular Pepper RNA showed more green fluorescent signals in the cytosol compared to EGFP-TetR. Significant change of protein localization in the case of EGFP-EZH2-tDeg or mCherry-NF-κB-tDeg with the circular Pepper RNA was not observed. It was concluded that tDeg is a versatile tag for RNA-dependent protein stabilization. All cells were stained with Hoechst dye. Scale bar, 40 μm. Normalized total cellular fluorescence (FIGS. **7**E, **7**F, and **7**G) of individual cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P.sub.EGFP-TetR4Deg=2.9×10.sup.−136; ****P.sub.EGFP-EZH2-tDeg1.1×10.sup.−120, ****P.sub.mCherry-NF-κB-tDeg=3.5×10.sup.−119 by unpaired two-tailed Student's t-test.

(8) FIGS. **8**A-**8**B demonstrate the optimization of a concatenated Pepper tag to image mRNAs in live cells. Pepper RNA-regulated fluorescent proteins were used to fluorescently tag mRNAs in live cells. As a first step, the best way to incorporate the Pepper aptamers in the 3′UTR of a transcript of interest was determined. In these experiments, a fluorescent protein (mNeonGreen).sub.2-tDeg and an mCherry mRNA reporter (FIG. **8**A) containing 3′UTR tags comprising 10 or 20 concatenated Pepper aptamers with and without a folding scaffold, F30, were expressed respectively. In the case of the (Pepper).sub.20 and (F30-2×Pepper).sub.10 tags, mobile green fluorescent puncta in the cytosol were observed (FIG. **8**B). A signal to noise ratio was evident when the (F30-2×Pepper).sub.10 tag (signal to noise ratio=1.8) was used, compared to the (Pepper).sub.20 tag (signal to noise ratio=1.5). However, puncta were not readily detectable with either the (Pepper).sub.10 tag or the (F30-1×Pepper).sub.10 tag. Therefore, the (F30-2×Pepper).sub.10 tag was used to image mRNAs in the subsequent experiments. Scale bar, 20 μm. This experiment was performed three times with similar results.

(9) FIGS. **9**A-**9**D show the design of Pepper tags for imaging mRNA. Design and sequences of four Pepper tags used in FIG. **8**B: (Pepper).sub.10 (FIG. **9**A; SEQ ID NO: 119), (F30-1×Pepper).sub.10 (FIG. **9**B; SEQ ID NO: 120), (Pepper).sub.20 (FIG. **9**C; SEQ ID NO: 121), and (F30-1×Pepper).sub.10 (FIG. **9**D; SEQ ID NO: 122).

(10) FIGS. **10**A-**10**C demonstrate the optimization of the number of fluorescent mNeonGreen monomers in the fluorescent protein for imaging mRNA in live cells. In FIG. **8**B, it was observed that (F30-2×Pepper).sub.10 is the optimal tag for imaging mRNAs in live cells. To further optimize the system of using Pepper RNA-regulated fluorogenic protein to image mRNAs, it was determined whether increasing the number of fluorescent mNeonGreen could increase the fluorescence signal to background noise ratio of the mobile green fluorescent puncta. In these experiments, an mCherry mRNA reporter tagged with (F30-2×Pepper).sub.10 and tandem fluorescent mNeonGreen with 2, 3, or 4 copies were transiently expressed, respectively, in cells. Here, an increase of fluorescence intensity of the green fluorescent puncta as the number of tandem mNeonGreen increased from 2, 3, to 4 copies, respectively (FIG. **10**B) and (FIG. **10**C) was observed. mRNAs tagged with (F30-1×Pepper).sub.10 using the (mNeonGreen).sub.4-tDeg fluorescent fusion protein were also re-tested. It was shown that puncta were detectable, but not as pronounced as when the (F30-2×Pepper).sub.10 tag was used. Thus, it was concluded that (mNeonGreen).sub.4-tDeg provides a high signal to noise ratio for imaging mRNAs. Scale bar, 20 μm. FIG. **10**C is a graph showing the fluorescence intensity of green fluorescent puncta of

individual cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P.sub.(Pepper)20:(F30-2×Pepper)10=4.6×10.sup.19; ****P.sub.(mNeonGreen)2-tDeg:(mNeonGreen)3-tDeg=7.7×10.sup.−9; ****P.sub.(mNeonGreen)2-tDeg:(mNeonGreen)4-tDeg=2.5×10.sup.−29; ****P.sub.(mNeonGreen)3-tDeg:(mNeonGreen)4-tDeg=2.0×10.sup.−9; ****P.sub.(F30-2×Pepper)10:(F30-1×Pepper)10=5.6×10.sup.−17 by one-way ANOVA.

(11) FIGS. **11**A-**11**C demonstrate that Pepper tag enables visualization of both nuclear and cytosolic mRNAs. FIG. **11**A is a schematic representation of the DNA plasmid constructs used for imaging mRNAs in the nucleus and cytosol. To image nascent transcription of mRNA, cells coexpressing an mCherry mRNA reporter containing a 3′UTR green Pepper mRNA tag, (F30-2×Pepper).sub.10, and a green fluorescent fusion protein, (mNeonGreen).sub.4-tDeg were imaged (FIG. **11**B). Cytosolic green fluorescent puncta reflecting mCherry mRNA transcripts and nuclear green fluorescent puncta, potentially reflecting mCherry mRNA transcripts were observed. Less green fluorescent puncta in the nucleus were observed as compared to the cytosol. This potentially reflects that most of the nuclear mCherry mRNA transcripts were exported out of the nucleus. Scale bar, 20 μm. FIG. **11**C is a graph providing summary data of cytosolic and nuclear mRNA fluorescence intensity in FIG. **11**B (n=201 fluorescent puncta). Values are means±s.d. This experiment was performed three times with similar results.

(12) FIGS. **12**A-**12**D demonstrate that Pepper tag and fluorescent fusion protein enable visualization of individual mRNAs. To examine whether the puncta observed when imaging Pepper-tagged mRNAs might be stable degradation intermediates, northern blot was performed on total RNA extracted from cells expressing (F30-2×Pepper).sub.10-tagged mCherry RNA transcripts with and without coexpressing the fluorescent fusion protein, (mNeonGreen).sub.4-tDeg. In these experiments, only full-length mRNA transcript was detected (FIG. **12**A). Therefore, it was concluded that the fluorescent puncta in cells largely reflects the full-length transcript, and that degraded or liberated Pepper aptamers do not accumulate in cells. To assess whether the mobile green fluorescent puncta seen in cells expressing Pepper-tagged mRNA represent single mRNAs, a previously described mRNA imaging method in which the resulting puncta were validated to represent single mRNA was used (Yan et al., “Dynamics of Translation of Single mRNA Molecules In Vivo,” *Cell* 165:976-89 (2016), which is hereby incorporated by reference in its entirety). This system uses 24 PP7 RNA hairpins in the 3′UTR of a reporter mRNA, and a 3×mCherry-CAAX protein fused to PCP (PP7 coat protein), the PP7-binding protein. The PCP-3×mCherry-CAAX fusion protein is anchored to the membrane via the CAAX sequence, which reduces puncta motility and facilitates quantitative fluorescence measurements. A PP7-containing reporter mRNA was imaged with and without the (F30-2×Pepper).sub.10 tag (FIG. **12**B). The (mNeonGreen).sub.4-tDeg fluorescent fusion protein was used to image the Pepper-tagged mRNAs. If the Pepper tag or the green fluorescent fusion protein caused mRNA to aggregate, the Pepper-tagged reporter mRNA puncta would have been expected to have higher red fluorescence (from PCP-3×mCherry-CAAX) compared to the reporter mRNA puncta without the Pepper tag. The results of these experiments showed that the red fluorescence intensity distribution of the reporter mRNA is not significantly different with and without the Pepper tag (FIG. **12**C) (Black bars, 19 cells, 485 mRNAs; Shaded bars, 13 cells, 384 mRNAs). This suggests that the Pepper tag and the green fluorescent fusion protein do not cause mRNA aggregation. Furthermore, colocalization between the green and magenta fluorescent puncta was observed only when the reporter mRNA contained the Pepper tag (FIG. **12**D). These results suggest that the green fluorescent puncta observed using the Pepper tag and green fluorescent fusion protein are indeed individual mRNAs. Scale bar, 5 μm (left panel in FIG. **12**D), 1 μm (right panel in FIG. **12**D). In FIG. **12**D, the experiment of reporter mRNA with Pepper was performed three times with similar results, the experiment of reporter mRNA without Pepper was performed twice with similar results.

(13) FIGS. **13**A-**13**E demonstrate that Pepper tag and fluorescent fusion protein do not have observable effects on mRNA turnover kinetics, mRNA translation efficiency, or proteasome

activity in cells. To test whether adding the Pepper tag to an mRNA transcript affects its stability, reporter plasmids expressing mCherry transcripts with and without the (F30-2×Pepper).sub.10 tag were constructed. HEK293T cells were transfected with these two reporter plasmids, respectively. In each case, the same cells were cotransfected with the (mNeonGreen).sub.4-tDeg fluorescent fusion protein. The cells were treated with 5 μg/mL actinomycin D to inhibit new transcription. The amount of reporter mRNA transcripts remaining at each time point was quantified by RT-qPCR at t=0, 1, 2, 4, and 6 hours of actinomycin D treatment. The results showed that fusing the Pepper tag to the reporter mRNA (half-life=5.9 hours) does not significantly affect its turnover rate compared to its untagged counterpart (half-life=6.0 hours) (FIG. **13**A). Thus, these data suggest that Pepper-tagged mRNA transcripts have similar turnover kinetics as mRNAs without the Pepper tag. Data were collected from 2 independent cell cultures. Values are means±s.d. To test whether adding the Pepper tag to an mRNA transcript affects its protein translation efficiency, the protein translation efficiency of an mCherry mRNA was compared with and without the (F30-2×Pepper).sub.10 Pepper tag. HEK293T cells expressing mCherry mRNA or mCherry-(F30-2×Pepper).sub.10 mRNA were harvested. The amount of mCherry protein and mCherry mRNA was quantified by western blotting and RT-qPCR, respectively. A slight decrease of mRNA levels in the Pepper-tagged mCherry mRNA was observed compared to its untagged counterpart (FIG. **13**C). The same phenomenon was also observed in the mCherry mRNA tagged with the 24×MS2 hairpins (Wu et al., "Synonymous Modification results in High-Fidelity Gene Expression of Repetitive Protein and Nucleotide Sequences," *Genes Dev.* 29:876-86 (2015), which is hereby incorporated by reference in its entirety). This may due to the longer transcript length associate with 3′UTR-tagged mRNAs. Protein translation efficiency was calculated by normalizing the amount of mCherry protein to the amount of mCherry mRNA (FIGS. **13**B-**13**D). No significant difference in protein translation efficiency was found between the untagged mCherry mRNA transcript and the Pepper-tagged mCherry mRNA transcript (FIG. **13**D). These results suggest that Pepper tag does not significantly affect protein translation of these mRNA reporter transcripts. Data were collected from 2 independent cell cultures. Values are means±s.d. Since the degradation mechanism of the fluorescent RNA-regulated fusion proteins described herein relies on ubiquitination and subsequent proteasomal degradation, expression of fluorescent RNA-regulated fusion proteins could lead to the overload of proteasome activity in cells. To test whether the expression of fluorescent RNA-regulated fusion proteins overloads proteasome activity, a RNA-regulated fluorescent fusion protein, (mNeonGreen).sub.4-tDeg was expressed in HEK293T cells. If the expression of (mNeonGreen).sub.4-tDeg overloads the activity of the proteasome, an accumulation of the ubiquitinated protein in cells would be expected. FIG. **13**E shows western blotting results using an anti-ubiquitin antibody of untransfected cells and cells expressing (mNeonGreen).sub.4-tDeg. Significant difference in the ubiquitinated proteins were not observed. As a control, untransfected cells treated with a proteasome inhibitor (10 μM MG132) for 5 hours showed a significant increase of the ubiquitinated proteins (FIG. **13**E). Thus, these results suggest that expression of fluorescent RNA-regulated fusion proteins does not overload proteasome activity in cells. Data shown here is a representative image from 2 independent cell cultures.

(14) FIGS. **14**A-**14**D demonstrate that Pepper tag does not disrupt the localization of mRNAs. To determine whether the Pepper tag disrupts an mRNA's proper cellular localization, an ER-targeting reporter mRNA was chosen, and its localization in cells was imaged using the (F30-2×Pepper).sub.10 Pepper tag and the (mNeonGreen).sub.4-tDeg fluorescent fusion protein (FIG. **14**A). This ER-targeting reporter mRNA encodes the first 29 amino acids of cytochrome p450, CytERM, and the encoding sequence of mCherry followed by (F30-2×Pepper).sub.10 in the 3′UTR (FIG. **14**A). During protein translation, the CytERM peptide will direct this reporter mRNA to the outer ER membrane, and confine the mRNA's mobility. Indeed, green fluorescent puncta with low mobility were observed (FIGS. **14**B, **14**D), suggesting that the reporter mRNA is localized to the outer ER membrane. To further validate the localization of the ER-targeting reporter mRNA, the

cells were treated with a translation inhibitor (100 μg/mL, puromycin) to liberate the reporter mRNA from the ER into the cytosol. A significant mobility increase of the green fluorescent puncta was observed (FIG. **14**C, FIG. **14**D), reflecting the dissociation of the reporter mRNA from the ER. Together, these results confirmed that the Pepper tag does not disrupt the localization of mRNAs. Scale bar in (FIG. **14**B, FIG. **14**C), 10 μm. Relative diffusion coefficient of mRNA puncta is plotted (n=2 independent cell cultures). Values are means±s.d. ****P=2.7×10.sup.−6 by unpaired two-tailed Student's t-test.

(15) FIGS. **15**A-**15**C demonstrate the imaging of green Pepper-tagged β-actin mRNA in live cells. FIG. **15**A shows DNA plasmid constructs used for imaging β-actin mRNA in live cells. To image β-actin mRNA localization in response to arsenite stress, a β-actin mRNA reporter containing a 3′UTR green Pepper mRNA tag, (F30-2×Pepper).sub.10 was constructed (FIG. **15**B). Cells coexpressing this β-actin mRNA reporter and a green fluorescent RNA-regulated fusion protein, (mNeonGreen).sub.4-tDeg were imaged before and 45 minutes after arsenite (500 μM) treatment to induce stress granules. Individual mRNA transcripts were observed to rapidly accumulated to form stress granules as evidenced by coexpression of tetramethylrhodamine-labeled HaloTag-G3BP1 to label stress granules. Scale bar, 20 μm. FIG. **15**C shows the fluorescence ratio of foci/cytosol in untreated cells vs. arsenite treated cells is plotted (n=3 independent cell cultures). Values are means±s.d. ****P=2.5×10.sup.−31 by unpaired two-tailed Student's t-test.

(16) FIGS. **16**A-**16**B demonstrate that (mNeonGreen).sub.4-tDeg without the Pepper-tagged β-actin mRNA does not accumulate in stress granules upon arsenite treatment. In FIGS. **15**A-**15**C, cytosolic green fluorescent puncta were shown to accumulate in stress granules to form foci upon application of 500 μM arsenite. However, the formation of green fluorescent foci in stress granules could be due to aggregation of the fluorescent RNA-regulated fusion protein, (mNeonGreen).sub.4-tDeg, regardless of the present of the β-actin mRNA. To test whether this is the case, (mNeonGreen).sub.4-tDeg was coexpressed with circular Pepper RNA in U2OS cells (FIG. **16**A). Before arsenite treatment, cytosolic green fluorescent was observed without any puncta, which is consistent with the results in FIGS. **5**A-**5**G. Upon application of 500 μM arsenite, green fluorescent foci formation was not observed (FIG. **15**B). These results confirmed that the formation of green fluorescent foci in FIGS. **15**A-**15**C were indeed due to the β-actin mRNA. This experiment was performed twice with similar results. Scale bar, 20 μm.

(17) FIGS. **17**A-**17**B demonstrate imaging of mRNAs using Pepper RNA-regulated fluorescent fusion proteins with different hues. So far, mRNA imaging using the green Pepper RNA tag, comprising the Pepper aptamer and a Pepper-regulated fluorescent mNeonGreen fusion protein has been described herein. To further expand the color palette for mRNA imaging, (mVenus).sub.2-tDeg and (mCherry).sub.2-tDeg were expressed to generate yellow Pepper and red Pepper complexes on mRNA. In these experiments, (mVenus).sub.2-tDeg was used to image an mCherry mRNA reporter tagged with (F30-2×Pepper).sub.10 (FIG. **17**A), and (mCherry).sub.2-tDeg was used to image a β-actin mRNA reporter tagged with (F30-2×Pepper).sub.10 (FIG. **17**B), respectively. In both cases, mobile fluorescent puncta were observed in cells. This experiment was performed twice with similar results. Scale bar, 20 μm.

(18) FIGS. **18**A-**18**D demonstrate the use of the tDeg-Pepper system to selectively biotinylate RNA-binding protein. tDeg was first shown to confer Pepper RNA-dependent regulation of a biotin ligase, TurboID, and a peroxidase, APEX2. HEK293T cells transiently expressed EGFP-TurboID-tDeg (FIG. **18**A), and EGFP-APEX2-tDeg (FIG. **18**B), with and without the Pepper RNA aptamer, respectively. In each case, proteins were nearly undetectable unless coexpressed with the Pepper RNA. FIG. **18**C is a schematic showing that a selectively activated biotin ligase (TurboID-tDeg) specifically biotinylates an RNA-binding protein (CELF1) that bind to the RNA sequence of interest (EDEN15). FIG. **18** D shows that TurboID-tDeg enables selective biotinylation of CELF1, while minimizing nonspecific biotinylation of proteins that do not bind to the RNA of interest (EDEN15).

(19) FIG. **19** demonstrates that Tat-GG confers Pepper RNA-dependent Regulation. In these experiments, U2OS cells transiently expressed mNeonGreen-Tat-GG fusion protein with and without the circular Pepper RNA aptamer, respectively. mNeonGreen was nearly undetectable (left panels) unless coexpressed with circular Pepper RNA (right panels). All cells were stained with Hoechst dye. Scale bar, 20 μm.
(20) FIG. **20** demonstrate that HIV Tat-RRRG (SEQ ID NO: 127) confers HIV TAR RNA-dependent regulation. In these experiments, cells transiently expressed YFP-HIV Tat-RRRG fusion protein with and without the circular HIV TAR RNA aptamer, respectively. YFP was nearly undetectable (top left panel) unless coexpressed with circular HIV TAR RNA aptamer (right panel). Bottom panels show brightfield microscopy of cells transfected with EYFP-HIV Tat-RRRG in the absence (left panel) or presence (right panel) of circular HIV TAR RNA (SEQ ID NO: 128).
DETAILED DESCRIPTION
(21) A first aspect of the disclosure relates to a nucleic acid molecule encoding an RNA-regulated fusion protein. The nucleic acid molecule includes: a first nucleic acid sequence encoding a protein of interest and a second nucleic acid sequence encoding an RNA-regulated destabilization domain, where the second nucleic acid sequence is operably coupled to the first nucleic acid sequence.
(22) The terms protein and polypeptide are generally used interchangeably and refer to a single polypeptide chain. It will be appreciated that such polypeptide chains may bind to other polypeptides or proteins, or other molecules such as cofactors. The terms protein and polypeptide also refer to variants, mutants, biologically active fragments, modifications, analogs and/or derivatives of the polypeptides described herein. The term fusion protein refers to a protein that is comprised of two or more amino acid sequences, from two or more proteins or polypeptide sequences that are not found linked in nature and that are physically linked by a peptide bond.
(23) A protein of interest refers to a protein/polypeptide that is desired and/or being assessed. In other words, a protein of interest may be any protein. In some embodiments, the protein of interest is a protein that is the subject of research. In some embodiments, the protein of interest is known to be involved in a disease state, and is specifically targeted in treatment of the disease state.
(24) In some embodiments, the protein of interest is a fluorescent protein, a bioluminescent protein, an enzyme, or a transcriptional regulator.
(25) In some embodiments, the protein of interest is a florescent protein. As used herein, the term “fluorescent protein” refers to a protein or polypeptide which fluoresces, or emits light, when excited with appropriate electromagnetic radiation.
(26) Suitable fluorescent proteins include, without limitation, Green Fluorescent Protein, Enhanced Green Fluorescent Protein (EGFP), Enhanced Yellow Fluorescent Protein (EYFP), Venus, mVenus, Citrine, mCitrine, Cerulean, mCerulean, Orange Fluorescent Protein (OFP), mNeonGreen, moxNeonGreen, mCherry, mTagBFP, Venus, mVenus, mTurquoise, mScarlet, mWasabi, mOrange, and dTomato. Suitable fluorescent protein amino acid sequences are shown in Table 1 below.
(27) TABLE-US-00001 TABLE 1 Exemplary Fluorescent Protein Amino Acid Sequences Fluorescent SEQ Protein Amino Acid Sequence ID NO: Green MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFI 1 Fluorescent CTIGKLPVPWPTLVITFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQ Protein ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKL (GFP) EYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPI GDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDE LYK Enhanced MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 2 Green ICTIGKLPVPWPTLVTILTYGVQCFSRYPDHMKQHDFFKSAMPEGYV Fluorescent QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK Protein LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTP (EGFP) IGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK Enhanced MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 3 Yellow

ICTIGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYV Fluorescent
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK Protein
LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTP (EYFP)
IGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK Venus
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKL 4
ICTIGKLPVPWPTLVTTLGYGLQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mVenus
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKL 5
ICTIGKLPVPWPTLVTTLGYGLQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK Citrine
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 6
ICTIGKLPVPWPTLVTTFGYGLMCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mCitrine
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 7
ICTIGKLPVPWPTLVITFGYGLMCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK Cerulean
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 8
ICTIGKLPVPWPTLVTILTWGVQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNAISDNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTP
IGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mCerulean
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF 9
ICTIGKLPVPWPTLVTILTWGVQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNAISDNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTP
IGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK Orange
MNLSKNVSVSVYMKGNVNNHEFEYDGEGGGDPYTGKYSMKMTLRGQN 10 Fluorescent
CLPFSYDIITTAFQYGFRVFTKYPEGIVDYFKDSLPDAFQWNRRIVF Protein
EDGGVLNMSSDITYKDNVLHGDVWAVGVNFPPNGPVMKNEIVMEEPT (OFP)
EETFTPKNGVLVGFCPKAYLLKDGSYYYGNMTTFYRSKKSGQAPPGY
HFVKHRLVKINVGHGFKTVEQTEYATAHVSDLPK mNeon
MVSKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPNDGYEEL 11 Green
NLKSTKGDLQFSPWILVPHIGYGFHQYLPYPDGMSPFQAAMVDGSGY
QVHRTMQFEDGASLTVNYRYTYEGSHIKGEAQVKGTGFPADGPVMTN
SLTAADWCRSKKTYPNDKTIISTFKWSYTTGNGKRYRSTARTTYTFA
KPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAFTDVMGMDELY K moxNeon
MVSKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPNDGYEEL 12 Green
NLKSTKGDLQFSPWILVPHIGYGFHQYLPYPDGMSPFQAAMVDGSGY
QVHRTMQFEDGASLTVNYRYTYEGSHIKGEAQVKGTGFPADGPVMTN
SLTAADWSRSKKTYPNDKTIISTFKWSYTTGNGKRYRSTARTTYTFA
KPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAFTDVMGMDELY K mCherry
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQ 13

TAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPE
GFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVM
QKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKA
KKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELY K (GenBank Accession No. QEM23462.1, which is hereby incorporated by reference in its entirety) mTagBFP MVSKGEELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTMRI 14 KVVEGGPLPFAFDILATSFLYGSKTFINHTQGIPDFFKQSFPEGFTW
ERVTTYEDGGVLTATQDTSLQDGCLIYNVKIRGVNFTSNGPVMQKKT
LGWEAFTETLYPADGGLEGRNDMALKLVGGSHLIANAKTTYRSKKPA
KNLKMPGVYYVDYRLERIKEANNETYVEQHEVAVARYCDLPSKLGHK LN Venus
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKL 15
ICTIGKLPVPWPTLVTTLGYGLQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mVenus
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKL 16
ICTIGKLPVPWPTLVTTLGYGLQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTP
IGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mTurquoise
MVSKGEELFTGVVPILVELDGDVNGHKFsysGEGEGDATyGKLTLKF 17
ICTIGKLPVPWPTLVTILSWGVQCFARYPDHMKQHDFFKSAMPEGYV
QERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
LEYNYISDNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTP
IGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMD ELYK mScarlet
MVSKGEAVIKEFMRFKVHMEGSMNGHEFEIEGEGEGRPYEGTQTAKL 18
KVIKGGPLPFSWDILSPQFMYGSRAFTKHPADIPDYYKQSFPEGFKW
ERVMNFEDGGAVIVTQDTSLEDGTLIYKVKLRGINFPPDGPVMQKKT
MGWEASTERLYPEDGVLKGDIKMALRLKDGGRYLADFKITYKAKKPV
QMPGAYNVDRKLDITSHNEDYTVVEQYERSEGRHSTGGMDELYK mWasabi
MVSKGEETTMGVIKPDMKIKLKMEGNVNGHAFVIEGEGEGKPYDGTN 19
TINLEVKEGAPLPFSYDILTTAFSYGNRAFTKYPDDIPNYFKQSFPE
GYSWERTMTFEDKGIVKVKSDISMEEDSFIYEIHLKGENFPPNGPVM
QKETTGWDASTERMYVRDGVLKGDVKMKLLLEGGGHHRVDFKTIYRA
KKAVKLPDYHFVDHRIEILNHDKDYNKVIVYETAVARNSTDGMDELY K mOrange
MVSKGEENNMAIIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGFQ 20
TAKLKVTKGGPLPFAWDILSPQFTYGSKAYVKHPADIPDYFKLSFPE
GFKWERVMNFEDGGVVIVTQDSSLQDGEFIYKVKLRGINFPSDGPVM
QKKTMGWEASSERMYPEDGALKGEIKMRLKLKDGGHYTSEVKITYKA
KKPVQLPGAYIVGIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELY K dTomato
MVSKGEEVIKEFMRFKVRMEGSMNGHEFEIEGEGEGRPYEGTQTAKL 21
KVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYKKLSFPEGFKW
ERVMNFEDGGLVTVTQDSSLQDGTLIYKVKMRGINFPPDGPVMQKKT
MGWEASTERLYPRDGVLKGEIHQALKLKDGGHYLVEFKTIYMAKKPV
QLPGYYYVDTKLDITSHNEDYTIVEQYERSEGRHHLFLYGMDELYK
(28) In other embodiments, the protein of interest is a bioluminescent protein. As used herein, the term “bioluminescent protein” refers to any protein capable of acting on a suitable substrate and producing luminescence. As used herein, the term “substrate” refers to any molecule capable of producing or absorbing luminescence with a bioluminescent protein. Suitable bioluminescent proteins include, without limitation, luciferase, β-galactosidase, β-lactamase, peroxidase, alkaline

phosphatase, β-glucuronidase, and β-glucosidase. Exemplary bioluminescent amino acid sequences are shown in Table 2 below.

(29) TABLE-US-00002 TABLE 2 Exemplary Bioluminescent Protein Amino Acid Sequences Bioluminescent SEQ ID Protein Amino Acid Sequence NO: Nanoluc MVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTP 22 luciferase IQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVYP (Nluc) VDDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKK ITVTGTLWNGNKIIDERLINPDGSLLFRVTINGVTGWRLCER ILA (GenBank Accession No. AFI79290.1, which is hereby incorporated by reference in its entirety) Firefly MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAF 23 luciferase TDAHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSEN SLQFFMPVLGALFIGVAVAPANDIYNERELLNSMNISQPTVV FVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFV TSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVAL PHRTACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTT LGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSF FAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGI RQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDT GKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHS GDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAELESILLQHP NIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVAS QVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGG KSKL (GenBank Accession No. CAB91857.1, which is hereby incorporated by reference in its entirety) Renilla MASKVYDPEQRKRMITGPQWWARCKQMNVLDSFINYYDSEKH 24 luciferase AENAVIFLHGNAASSYLWRHVVPHIEPVARCIIPDLIGMGKS (Rluc) GKSGNGSYRLLDHYKYLTAWFELLNLPKKIIFVGHDWGACLA FHYSYEHQDKIKAIVHAESVVDVIESWDEWPDIEEDIALIKS EEGEKMVLENNFFVETMLPSKIMRKLEPEEFAAYLEPFKEKG EVRRPTLSWPREIPLVKGGKPDVVQIVRNYNAYLRASDDLPK MFIESDPGFFSNAIVEGAKKFPNTEFVKVKGLHFSQEDAPDE MGKYIKSFVERVLKNEQ (GenBank Accession No. ABA41680.1, which is hereby incorporated by reference in its entirety) Gaussia MGVKVLFALICIAVAEAKPTENNEDFNIVAVASNFATTDLDA 25 luciferase DRGKLPGKKLPLEVLKEMEANARKAGCTRGCLICLSHIKCTP KMKKFIPGRCHTYEGDKESAQGGIGEAIVDIPEIPGFKDLEP MEQFIAQVDLCVDCTTGCLKGLANVQCSDLLKKWLPQRCATF ASKIQGQVDKIKGAGGD (GenBank Accession No. BAR71165.1, which is hereby incorporated by reference in its entirety) β-galactosidase VVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQ 26 LRSLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQM HGYDAPIYTNVTYPITVNPPFVPTENPTGCYSLTFNVDESWL QEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAF LRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHKP TTQISDFHVATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSL WQGETQVASGTAPFGGEIIDERGGYADRVTLRLNVENPKLWS AEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLLLL NGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNA VRCSHYPNHPLWYTLCDRYGLYVVDEANIETHGMVPMNRLTD DPRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHDA LYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPF

PAVPKWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQ
AFRQYPRLQGGFVWDWVDQSLIKYDENGNPWSAYGGDFGDTP
NDRQFCMNGLVFADRTPHPALIEAKHQQQFFQFRLSGQTIEV
TSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLI
ELPELPQPESAGQLWLTVRVVQPNATAWSEAGHISAWQQWRL
AENLSVTLPAASHAIPHLTTSEMDFCIELGNKRWQFNRQSGF
LSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRIDPNAW
VERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKT
LFISRKTYRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQ
VAERVNWLGLGPQENYPDRLTAACFDRWDLPLSDMYTPYVFP
SENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETSHRH
LLHAEEGTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHY QLVWCQK (GenBank Accession No. CAB90353.1, which is hereby incorporated by reference in its entirety) β-lactamase MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARV 27 (HaloTag) GYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDA
GQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAAIT
MSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELN
EAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWME
ADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGK
PSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW (GenBank Accession No. AEQ28652.1, which is hereby incorporated by reference in its entirety) Ascorbate MGKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLMLRLAWH 28 peroxidase 1, SAGTFDKGTKTGGPFGTIKHPAELAHSANNGLDIAVRLLEPL cytosolic
KAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKPEPP (*Glycine max*)
PEGRLPDATKGSDHLRDVFGKAMGLTDQDIVALSGGHTIGAA
HKERSGFEGPWTSNPLIFDNSYFTELLSGEKEGLLQLPSDKA
LLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFADA (GenBank Accession No. NP_001237785.1, which is hereby incorporated by reference in its entirety) Ascorbate MTKNYPTVSEDYKKAVEKCRRKLRGLIAEKNCAPIMVRLAWH 29 peroxidase 1
SAGTFDCQSRTGGPFGTMRFDAEQAHGANSGIHIALRLLDPI (*Arabidopsis*
REQFPTISFADFHQLAGVVAVEVTGGPDIPFHPGREDKPQPP *thaliana*)
PEGRLPDATKGCDHLRDVFAKQMGLSDKDIVALSGAHTLGRC
HKDRSGFEGAWTSNPLIFDNSYFKELLSGEKEGLLQLVSDKA
LLDDPVFRPLVEKYAADEDAFFADYAEAHMKLSELGFADA (GenBank Accession No. NP_172267.1, which is hereby incorporated by reference in its entirety) Ascorbate MVKKSYPEVKEEYKKAVQRCKRKLRGLIAEKHCAPIVLRLAW 30 peroxidase 2
HSAGTFDVKTKTGGPFGTIRHPQELAHDANNGLDIAVRLLDP (*Arabidopsis*
IKELFPILSYADFYQLAGVVAVEITGGPEIPFHPGRLDKVEP *thaliana*)
PPEGRLPQATKGVDHLRDVFGRMGLNDKDIVALSGGHTLGRC
HKERSGFEGAWTPNPLIFDNSYFKEILSGEKEGLLQLPTDKA
LLDDPLFLPFVEKYAADEDAFFEDYTEAHLKLSELGFADKE (GenBank Accession No. AEE74792.1, which is hereby incorporated by reference in its entirety) Ascorbate MGKSYPTVSPDYQKAIEKAKRKLRGFIAEKKCAPLILRLAWH 31 peroxidase
SAGTFDSKTKTGGPFGTIKHQAELAHGANNGLDIAVRLLEPI (*Pisum sativum*)
KEQFPIVSYADFYQLAGVVAVEITGGPEVPFHPGREDKPEPP
PEGRLPDATKGSDHLRDVFGKAMGLSDQDIVALSGGHTIGAA
HKERSGFEGPWTSNPLIFDNSYFTELLTGEKDGLLQLPSDKA
LLTDSVFRPLVEKYAADEDVFFADYAEAHLKLSELGFAEA (GenBank Accession No. AAA33645. 1, which is hereby incorporated by reference in its entirety) APEX2 MGKSYPTVSADYQDAVEKAKKKLRGFIAEKRCAPLMLRLAFH 32 (soybean

SAGTFDKGTKTGGPFGTIKHPAELAHSANNGLDIAVRLLEPL ascorbate
KAEFPILSYADFYQLAGVVAVEVTGGPKVPFHPGREDKPEPP peroxidase)
PEGRLPDPTKGSDHLRDVFGKAMGLTDQDIVALSGGHTIGAA
HKERSGFEGPWTSNPLIFDNSYFTELLSGEKEGLLQLPSDKA
LLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFADA (see, e.g., Ganapathy et al., "Compartment-Specific Labeling of Bacterial Periplasmic Proteins by Peroxidase-Mediated Biotinylation," ACS Infect. Dis. 4(6): 918-925 (2018) and Lam et al., "Directed Evoluation of APEX2 for Electron Microscopy and Proximity Labeling," Nature Methods 12:51-54 (2014), which are hereby incorporated by reference in their entirety) Horseradish
MQLTPTFYDNSCPNVSNIVRDTIVNELRSDPRIAASILRLHF 33 peroxidase
HDCFVNGCDASILLDNTTNANSARGFPVIDRMKAAVESACPR (*Armoracia*
TVSCADLLTIAAQQSVTLAGGPSWRVPLGRRDSLQAFLDLAN *rusticana*)
ANLPAPFFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQC
RFIMDRLYNFSNTGLPDPILNITYLQTLRGLCPLNGNLSALV
DFDLRTPTIFDNKYYVNLEEQKGLIQSDQELFSSPNATDTIP
LVRSFANSTQTFFNAFVEAMDRMGNITPLTGTQGQIRLNCRV VNSNS (GenBank Accession No. CAA00083.1, which is hereby incorporated by reference in its entirety) Alkaline MKQSTIALALLPLLFTPVTKARTPEMPLQGTAVDGGGGSMHA 34
phosphatase SLEVLENRAAQGDITAPGGARRLTGDQTAALRDSLSDKPAKN
IILLIGDGMGDSEITAARNYAEGAGGFFKGIDALPLTGQYTH
YALNKKTGKPDYVTDSAASATAWSTGVKTYNGALGVDIHEKD
HPTILEMAKAAGLATGNVSTAELQDATPAALVAHVTSRKCYG
PSATSEKCPGNALEKGGKGSITEQLLNARADVTLGGGAKTFA
ETATAGEWQGKTLREQAQARGYQLVSDAASLNSVTEANQQKP
LLGLFADGNMPVRWLGPKATYHGNIDKPAVTCTPNPQRNDSV
PTLAQMTDKAIELLSKNEKGFFLQVEGASIDKQDHAANPCGQ
IGETVDLDEAVQRALEFAKKEGNTLVIVTADHAHASQIVAPD
TKAPGLTQALNTKDGAVMVMSYGNSEEDSQEHTGSQLRIAAY
GPHAANVVGLTDQTDLFYTMKAALGLK (GenBank Accession No. AAK73766.1, which is hereby incorporated by reference in its entirety) Alkaline
MKQSTIALALLPLLFTPVTKARTPEMPVLENRAAQGDITAPG 35 phosphatase
GARRLTGDQTAALRDSLSDKPAKNIILLIGDGMGDSEITAAR (*Escherichia*
NYAEGAGGFFKGIDALPLTGQYTHYALNKKTGKPDYVTDSAA *coli*)
SATAWSTGVKTYNGALGVDIHEKDHPTILEMAKAAGLATGNV
STAELQDATPAALVAHVTSRKCYGPSATSEKCPGNALEKGGK
GSITEQLLNARADVTLGGGAKTFAETATAGEWQGKTLREQAQ
ARGYQLVSDAASLNSVTEANQQKPLLGLFADGNMPVRWLGPK
ATYHGNIDKPAVTCTPNPQRNDSVPTLAQMTDKAIELLSKNE
KGFFLQVEGASIDKQDHAANPCGQIGETVDLDEAVQRALEFA
KKEGNTLVIVTADHAHASQVVAPDTKAPGLTQALNTKDGAVM
VMSYGNSEEDSQEHTGSQLRIAAYGPHAANVVGLTDQTDLFY TMKAALGLK (GenBank Accession No. WP_001364609.1, which is hereby incorporated by reference in its entirety) β-glucuronidase MLRPVETPTREIKKLDGLWAFSLDRENCGIDQRWWESALQES
36 (*Escherichia* RAIAVPGSFNDQFADADIRNYAGNVWYQREVFIPKGWAGQRI *coli*)
VLRFDAVTHYGKVWVNNQEVMEHQGGYTPFEADVTPYVIAGK
SVRITVCVNNELNWQTIPPGMVITDENGKKKQSYFHDFFNYA
GIHRSVMLYTTPNTWVDDITVVTHVAQDCNHASVDWQVVANG
DVSVELRDADQQVVATGQGTSGTLQVVNPHLWQPGEGYLYEL
CVTAKSQTECDIYPLRVGIRSVAVKGQQFLINHKPFYFTGFG

RHEDADLRGKGFDNVLMVHDHALMDWIGANSYRTSHYPYAEE
MLDWADEHGIVVIDETAAVGFNLSLGIGFEAGNKPKELYSEE
AVNGETQQAHLQAIKELIARDKNHPSVVMWSIANEPDTRPQV
HGNISPLAEATRKLDPTRPITCVNVMFCDAHTDTISDLFDVL
CLNRYYGWYVQSGDLETAEKVLEKELLAWQEKLHQPIIITEY
GVDTLAGLHSMYTDMWSEEYQCAWLDMYHRVFDRVSAVVGEQ
VWNFADFATSQGILRVGGNKKGIFTRDRKPKSAAFLLQKRWT GMNFGEKPQQGGKQ (GenBank Accession No. AAC53703.1, which is hereby incorporated by reference in its entirety) β-glucosidase
MSTNSNIRQKLGQLIMMDFRYWGEDSNNQRIPFTKINDIVNK 37 (*Francisella*
IFKDYNLGGFILFRENIQNNEQVISLLRDLQANTNTPIFFAT *tularensis*)
DQEGGRVNRLQQGTSGCGNMALAATDNPHNAYTMAKIIGDEL
YSLGININFAPAVDVNSNKNNPIIGVRSYSDNPDIVIDYAKN
AINGYHDAKIIDCIKHFPGHGDTATDSHLGNVNLDKTLKELQ
TTELLPFSKLARDCSMIMTAHISVPALDDTQYQSVSTSENIY
VPATLSYKIITKLLKQQMKFDGLVVSDAMDMHAIAKHFGTIE
ASKLAILAGIDILLMPVRVWSENDLYKLEELFCELEKGYNQN
SNFANAVDNVYTNITDFKAKHKLDESLIFKLSQDEQLKYANQ
IVNSNKHQQIALDIAKQSTTVVKNSGIIPCDLNKLKNILIVD
SDNQRLADFHSELQKIVLDNNSNVIINCENINNHNIKTIIEN
ADLILLISANLREYNQTYSYITSIKPEQTINIAALTPYDINY
IDNIINYVCIYGATSMDQTNYTKTSLKINIQTTLENIFGNKE IKGVLPVSL (GenBank Accession No. AAC53703.1, which is hereby incorporated by reference in its entirety)

(30) The protein of interest may be an enzyme. In some embodiments, the enzyme is selected from the group consisting of a ligase and a methyltransferase.

(31) As described herein, the term “ligase” refers to an enzyme that catalyzes the joining of two large molecules by forming a new chemical bond, usually with accompanying hydrolysis of a small pendant chemical group on one of the larger molecules or the enzyme catalyzing the linking together of two compounds. Suitable ligases include, without limitation, DNA ligases, RNA ligases, amino acid—tRNA ligases (e.g., tyrosine—tRNA ligase, tryptophan—tRNA ligase, threonine—tRNA ligase, leucine—tRNA ligase, isoleucine—tRNA ligase, lysine—tRNA ligase, alanine—tRNA ligase, valine—tRNA ligase, methionine—tRNA ligase, serine—tRNA ligase, aspartate—tRNA ligase, D-alanine—tRNA ligase, glycine—tRNA ligase, proline—tRNA ligase, cysteine—tRNA ligase, glutamate—tRNA ligase, glutamine—tRNA ligase, arginine—tRNA ligase, phenylalanine—tRNA ligase, histidine—tRNA ligase, asparagine—tRNA ligase, aspartate—tRNA ligase, glutamate—tRNA ligase), acetate—CoA ligase, succinate—CoA ligase, biotin—CoA ligase (i.e., biotin ligase), carboxylic acid—CoA ligase, acetate—CoA ligase, and aspartate—ammonia ligase (see, e.g., McDonald, Andrew, “The Enzyme List Class 6—Ligases,” *ExplorEnz Database* (2019), which is hereby incorporated by reference in its entirety).

(32) In some embodiments, the ligase is a biotin ligase. As described herein, biotin ligases catalyze the formation of biotin-5′-AMP anhydride, which diffuses out of the active site to biotinylate proximal endogenous proteins on nucleophilic residues such as lysine. In some embodiments, the biotin ligase is selected from TurboID, miniTurbo, and *E. coli* BirA (see, e.g., Branon et al., “Efficient Proximity Labeling in Living Cells and Organisms with TurboID,” *Nat. Biotechnol.* 36(9):880-887 (2018), which is hereby incorporated by reference in its entirety).

(33) The methyltransferase may be a histone methyltransferase, an N-terminal methyltransferase, a DNA/RNA methyltransferase, a natural product methyltransferase, a non-SAM dependent methyltransferase, or a radical SAM methyltransferase. As described herein, histone methyl transferases catalyze the transfer of one, two, or three methyl groups to lysine and arginine residues

of histone proteins. In some embodiments, the histone methyltransferase is a histone-lysine N-methyltransferase selected from the group consisting of enhancer of zeste homolog 1 (EZH1), enhancer of zeste homolog 2 (EZH2), disruptor of telomeric silencing 1-like (DOT1-like), ASH1L, euchromatic histone-lysine N-methyltransferase 1 (EHMT1), euchromatic histone-lysine N-methyltransferase 2 (EHMT2), histone-lysine N-methyltransferase 2A, histone-lysine N-methyltransferase 2D (KMT2D), lysine N-methyltransferase 2C (KMT2C), myeloid/lymphoid or mixed-lineage leukemia 4 (MLL4), lysine methyltransferase 2E, and nuclear receptor binding SET domain protein 1 (NSD1). In other embodiments, the histone methyltransferase is a histone-arginine N-methyltransferases selected from the group consisting of protein arginine N-methyltransferase 1, protein arginine N-methyltransferase 3, protein arginine N-methyltransferase 4, protein arginine N-methyltransferase 5, and protein arginine N-methyltransferase 7.

(34) Non-limiting examples of suitable enzymes are identified in Table 3 below.

(35) TABLE-US-00003 TABLE 3 Exemplary Enzyme Amino Acid Sequences SEQ ID Enzyme Amino Acid Sequence NO: *E. coli* BirA MKDNTVPLKLIALLANGEFHSGEQLGETLGMSRAAINKHIQTLR 38 (Biotin-CoA DWGVDVFTVPGKGYSLPEPIQLLNAKQILGQLDGGSVAVLPVID ligase) STNQYLLDRIGELKSGDACIAEYQQAGRGRRGRKWFSPFGANLY LSMFWRLEQGPAAAIGLSLVIGIVMAEVLRKLGADKVRVKWPND LYLQDRKLAGILVELTGKTGDAAQIVIGAGINMAMRRVEESVVN QGWITLQEAGINLDRNTLAAMLIRELRAALELFEQEGLAPYLSR WEKLDNFINRPVKLIIGDKEIFGISRGIDKQGALLLEQDGIIKP WMGGEISLRSAEK (GenBank Accession No. NP_418404.1, which is hereby incorporated by reference in its entirety) miniTurbo MIPLLNAKQILGQLDGGSVAVLPVVDSTNQYLLDRIGELKSGDA 39 biotin ligase CIAEYQQAGRGSRGRKWFSPFGANLYLSMFWRLKRGPAAIGLGP VIGIVMAEALRKLGADKVRVKWPNDLYLQDRKLAGILVELAGIT GDAAQIVIGAGINVAMRRVEESVVNQGWITLQEAGINLDRNTLA AMLIRELRAALELFEQEGLAPYLSRWEKLDNFINRPVKLIIGDK EIFGISRGIDKQGALLLEQDGVIKPWMGGEISLRSAEK (see, e.g., Branon et al., “Efficient Proximity Labeling in Living Cells and Organisms with TurboID,” Nat. Biotechnol. 36(9):880-887 (2018), which is hereby incorporated by reference in its entirety) Turbo ID MKDNTVPLKLIALLANGEFHSGEQLGETLGMSRAAINKHIQTLR 40 biotin ligase DWGVDVFTVPGKGYSLPEPIPLLNAKQILGQLDGGSVAVLPVVD STNQYLLDRIGELKSGDACIAEYQQAGRGSRGRKWFSPFGANLY LSMFWRLKRGPAAIGLGPVIGIVMAEALRKLGADKVRVKWPNDL YLQDRKLAGILVELAGITGDAAQIVIGAGINVAMRRVEESVVNQ GWITLQEAGINLDRNTLAATLIRELRAALELFEQEGLAPYLPRW EKLDNFINRPVKLIIGDKEIFGISRGIDKQGALLLEQDGVIKPW MGGEISLRSAEK (see, e.g., Branon et al., “Efficient Proximity Labeling in Living Cells and Organisms with TurboID,” Nat. Biotechnol. 36(9):880-887 (2018), which is hereby incorporated by reference in its entirety) Biotin ligase MDYKDDDDKSPRSMKDNTVPLKLIALLANGEFHSGEQLGETLGM 41 (Mammalian SRAAINKHIQTLRDWGVDVFTVPGKGYSLPEPIQLLNAKQILGQ expression LDGGSVAVLPVIDSTNQYLLDRIGELKSGDACIAEYQQAGRGRR vector GRKWFSPFGANLYLSMFWRLEQGPAAAIGLSLVIGIVMAEVLRK pCBio) LGADKVRVKWPNDLYLQDRKLAGILVELTGKTGDAAQIVIGAGI NMAMRRVEESVVNQGWITLQEAGINLDRNTLAAMLIRELRAALE LFEQEGLAPYLSRWEKLDNFINRPVKLIIGDKEIFGISRGIDKQ GALLLEQDGIIKPWMGGEISLRSAEK (GenBank Accession No. ABF74577.1, which

is hereby incorporated by reference in its entirety) Enhancer of
MGQTGKKSEKGPVCWRKRVKSEYMRLRQLKRFRRADEVKSMFSS 42 Zeste
NRQKILERTEILNQEWKQRRIQPVHILTSVSSLRGTRECSVISD Homolog 2
LDFPTQVIPLKTLNAVASVPIMYSWSPLQQNFMVEDETVLHNIP (*Homo*
YMGDEVLDQDGTFIEELIKNYDGKVHGDRECGFINDEIFVELVN *sapiens*)
ALGQYNDDDDDDDGDDPEEREEKQKDLEDHRDDKESRPPRKFPS methyl-
DKIFEAISSMFPDKGTAEELKEKYKELTEQQLPGALPPECTPNI transferase
DGPNAKSVQREQSLHSFHTLFCRRCFKYDCFLHPFHATPNTYKR
KNTETALDNKPCGPQCYQHLEGAKEFAAALTAERIKTPPKRPGG
RRRGRLPNNSSRPSTPTINVLESKDTDSDREAGTETGGENNDKE
EEEKKDETSSSSEANSRCQTPIKMKPNIEPPENVEWSGAEASMF
RVLIGTYYDNFCAIARLIGTKTCRQVYEFRVKESSIIAPAPAED
VDTPPRKKKRKHRLWAAHCRKIQLKKDGSSNHVYNYQPCDHPRQ
PCDSSCPCVIAQNFCEKFCQCSSECQNRFPGCRCKAQCNTKQCP
CYLAVRECDPDLCLICGAADHWDSKNVSCKNCSIQRGSKKHLLL
APSDVAGWGIFIKDPVQKNEFISEYCGEIISQDEADRRGKVYDK
YMCSFLFNLNNDFVVDATRKGNKIRFANHSVNPNCYAKVMMVNG
DHRIGIFAKRAIQTGEELFFDYRYSQADALKYVGIEREMEIP (GenBank Accession No. AAC51520.1, which is hereby incorporated by reference in its entirety)

(36) Additional suitable proteins of interest include, but are not limited to, a G-protein coupled receptor (GPCR), a nuclear receptor, a voltage gated ion channel, a ligand gated channel, a receptor tyrosine kinase, a growth factor, a phosphatase, a protein kinase, a viral regulator, a bacterial cell division protein, a scaffold protein, a DNA repair protein, a cytoskeletal protein, a ribosome, a histone deacetylase, an apoptosis regulator, a chaperone protein, a kinase, a phosphorylase, a phosphatase, deacetylase, a cytoskeletal protein (e.g., myosin, actin, dynein, kinesin, and tubulin).

(37) As described herein, a G-protein coupled receptor (GPCR) refers to a membrane protein which binds to a signaling molecule. Upon binding, a conformational change occurs, which allows binding of the GPCR to, and activation of, a G-protein. The activated G-protein then interacts with an effector molecule, which is typically involved in a second messenger pathway. Suitable G-protein coupled receptors may be selected from the group consisting of a luteinizing hormone receptor, a follicle stimulating hormone receptor, a thyroid stimulating hormone receptor, a calcitonin receptor, a glucagon receptor, a glucagon-like peptide 1 receptor (GLP-1), a metabotropic glutamate receptor, a parathyroid hormone receptor, a vasoactive intestinal peptide receptor, a secretin receptor, a growth hormone releasing factor (GRF) receptor, protease-activated receptors (PARs), cholecystokinin receptors, somatostatin receptors, melanocortin receptors, nucleotide receptors (e.g., ADP receptors), adenosine receptors, thromboxane receptors, platelet activating factor receptor, adrenergic receptors, 5-hydroxytryptamine (5-HT) receptors, a chemokine receptor (e.g., CXCR4, CCR5), chemokine receptors, neuropeptide receptors, opioid receptors, erythropoietin receptor, von Willebrand receptor, parathyroid hormone (PTH) receptor, vasoactive intestinal peptide (VIP) receptor, and collagen receptors. Exemplary protease-activated receptors include, without limitation, PAR1, PAR2, PAR3, or PAR4 receptors.

(38) In some embodiments, the protein of interest is a transcription factor. Transcription factors include proteins that are involved in gene regulation in prokaryotic and/or eukaryotic organisms. In one embodiment, transcription factors have a positive effect on gene expression and, thus, may be referred to as an activator or a transcriptional activation factor. In another embodiment, a transcription factor negatively regulates gene expression and, thus, may be referred to as a repressor or a transcription repression factor. Suitable transcription factors include, without limitation, c-Myc, c-Fos, c-Jun, CREB, GATA-2, GAL4, GAL4Np16, c-Myb, MyoD, and NFκB, and tetR. Exemplary transcription factors are identified in Table 4 below.

(39) TABLE-US-00004 TABLE 4 Exemplary Transcription Factor Amino Acid

Sequences Transcription SEQ ID Factor Amino Acid Sequence NO: c-Myc
MPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAP 43 (*Homo sapiens*)
SEDIWKKFELLPTPPLSPSRRSGLCSPSYVAVTPFSLRGDNDGGG
GSFSTADQLEMVTELLGGDMVNQSFICDPDDETFIKNIIIQDCMW
SGFSAAAKLVSEKLASYQAARKDSGSPNPARGHSVCSTSSLYLQD
LSAAASECIDPSVVFPYPLNDSSSPKSCASQDSSAFSPSSDSLLS
STESSPQGSPEPLVLHEETPPTTSSDSEEEQEDEEEIDVVSVEKR
QAPGKRSESGSPSAGGHSKPPHSPLVLKRCHVSTHQHNYAAPPST
RKDYPAAKRVKLDSVRVLRQISNNRKCTSPRSSDTEENVKRRTHN
VLERQRRNELKRSFFALRDQIPELENNEKAPKVVILKKATAYILS
VQAEEQKLISEEDLLRKRREQLKHKLEQLRNSCA (GenBank Accession No. AAA36340.1, which is hereby incorporated by reference in its entirety) c-Fos
MMFSGFNADYEASSSRCSSASPAGDSLSYYHSPADSFSSMGSPVN 44 (*Homo sapiens*)
AQDFCTDLAVSSANFIPTVTAISTSPDLQWLVQPALVSSVAPSQT
RAPHPFGVPAPSAGAYSRAGVVKTMTGGRAQSIGRRGKVEQLSPE
EEEKRRIRRERNKMAAAKCRNRRRELTDTLQAETDQLEDEKSALQ
TEIANLLKEKEKLEFILAAHRPACKIPDDLGFPEEMSVASLDLTG
GLPEVATPESEEAFTLPLLNDPEPKPSVEPVKSISSMELKTEPFD
DFLFPASSRPSGSETARSVPDMDLSGSFYAADWEPLHSGSLGMGP
MATELEPLCIPVVICTPSCTAYTSSFVFTYPEADSFPSCAAAHRK
GSSSNEPSSDSLSSPTLLAL (GenBank Accession No. AAA52471.1, which is hereby incorporated by reference in its entirety) c-Jun
MTAKMETTFYDDALNASFLPSESGPYGYSNPKILKQSMTLNLADP 45 (*Homo sapiens*)
VGSLKPHLRAKNSDLLTSPDVGLLKLASPELERLIIQSSNGHITT
TPTPTQFLCPKNVTDEQEGFAEGFVRALAELHSQNTLPSVTSAAQ
PVNGAGMVAPAVASVAGGSGSGGFSASLHSEPPVYANLSNFNPGA
LSSGGGAPSYGAAGLAFPAQPQQQQQPPHHLPQQMPVQHPRLQAL
KEEPQTVPEMPGETPPLSPIDMESQERIKAERKRMRNRIAASKCR
KRKLERIARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVMNH
VNSGCQLMLTQQLQTF (GenBank Accession No. NP_002219.1, which is hereby incorporated by reference in its entirety) CREB
MIMESGAENQQSGDAAVTEAENQQMTVQAQPQIATLAQVSMPAAH 46 (*Homo sapiens*)
ATSSAPTVTLVQLPNGQTVQVHGVIQAAQPSVIQSPQVQTVQIST
IAESEDSQESVDSVTDSQKRREILSRRPSYRKILNDLSSDAPGVP
RIEEEKSEEETSAPAITIVTVPTPIYQTSSGQYIAITQGGAIQLA
NNGTDGVQGLQTLTMTNAAATQPGTTILQYAQTTDGQQILVPSNQ
VVVQAASGDVQTYQIRTAPTSTIAPGVVMASSPALPTQPAEEAAR
KREVRLMKNREAARECRRKKKEYVKCLENRVAVLENQNKTLIEEL KALKDLYCHKSD (GenBank Accession No. AAA35715. 1, which is hereby incorporated by reference in its entirety) GATA-2
MEVAPEQPGWMAHPAVLNAQHPDSHHPGLAHNYMEPAHVLPPDEV 47 (*Homo sapiens*)
DVFFNHLDSQGNPYYANPAQRGVSYSPAHARLTGGQMCRPHLLHS
PGLPWLDGGKAALSAAHHKTWTVSPFSKTPLHPSAAGGPGGHSLC
TQGLGVGGGSSGSSVASLTPTAAHSGSHLFGFPPRHPKELSPDPS
TTGAASPASSSAGGSSARGEDKDGVKYQASLTESMKMESGRPLRP
GLATMGTQPATHHPIPTYPSYVPAAAHDYSSGLFHPGSFLGGPAS
SFTPKQRSKTRSCSEGRECVNCGATATPLWRRDGTGHYLCNACGF
YHKMKGQNRPLIKPKRRLSAARRAGTCCANCQTTITTLWRRNANG
DPVCNACGLYYKLHNVNRPLTMKKEGIQTRNRKMSNKSKKSKKGA
ECFEELSKCMQEKSSPFSAAALAGHMAPMGHLPPFSHSGHILPTP

TPIHPSSSLSFGHPHPSSMVTAMG (GenBank Accession No. AAA35869. 1, which is hereby incorporated by reference in its entirety) GAL4 MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTK 48 (*Saccharomyce* RSPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIK *revisiae*) ALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSEE SSNKGQRQLTVSIDSAAHHDNSTIPLDFMPRDALHGFDWSEEDDM SDGLPFLKTDPNNNGFFGDGSLLCILRSIGFKPENYTNSNVNRLP TMITDRYTLASRSTTSRLLQSYLNNFHPYCPIVHSPTLMMLYNNQ IEIASKDQWQILFNCILAIGAWCIEGESTDIDVFYYQNAKSHLTS KVFESGSIILVTALHLLSRYTQWRQKTNTSYNFHSFSIRMAISLG LNRDLPSSFSDSSILEQRRRIWWSVYSWEIQLSLLYGRSIQLSQN TISFPSSVDDVQRTTTGPTIYHGIIETARLLQVFTKIYELDKTVT AEKSPICAKKCLMICNEIEEVSRQAPKFLQMDISTTALTNLLKEH PWLSFTRFELKWKQLSLIIYVLRDFFTNFTQKKSQLEQDQNDHQS YEVKRCSIMLSDAAQRTVMSVSSYMDNHNVTPYFAWNCSYYLFNA VLVPIKTLLSNSKSNAENNETAQLLQQINTVLMLLKKLATFKIQT CEKYIQVLEEVCAPFLLSQCAIPLPHISYNNSNGSAIKNIVGSAT IAQYPTLPEENVNNISVKYVSPGSVGPSPVPLKSGASFSDLVKLL SNRPPSRNSPVTIPRSTPSHRSVTPFLGQQQQLQSLVPLTPSALF GGANFNQSGNIADSSLSFTFTNSSNGPNLITTQTNSQALSQPIAS SNVHDNFMNNEITASKIDDGNNSKPLSPGWTDQTAYNAFGITTGM FNTTTMDDVYNYLFDDEDTPPNPKKE (GenBank Accession No. AAA34626. 1, which is hereby incorporated by reference in its entirety) GAL4Np16 MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTK 49 (*Saccharomyce* RSPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIK *revisiae*) ALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSEE SSNKGQRQLTVSIEFSRGRTRNNYGSTIEGLLDLPDDDDAPAEAG LVAPRMSFLSAGQRPRRLSTTAPITDVSLVDELRLDGEEVDMTPA DALDDFDLEMLGDVESPSPGMTHDPVSYGALDVDDFEFEQMFTDA LGIDDFGG (GenBank Accession No. AAN86074.1, which is hereby incorporated by reference in its entirety) c-Myb MARRPRHSIYSSDEDDEDFEMCDHDYDGLLPKSGKRHLGKTRWTR 50 (*Homo sapiens*) EE (GenBank Accession No. AAA72118. 1, which is hereby incorporated by reference in its entirety) MyoD MELLSPPLRDIDLTGPDGSLCSFETADDFYDDPCFDSPDLRFFED 51 (*Mus musculus*) LDPRLVHVGALLKPEEHAHFSTAVHPGPGAREDEHVRAPSGHHQA GRCLLWACKACKRKTTNADRRKAATMRERRRLSKVNEAFETLKRC ISSNPNQRLPKVEILRNAIRYIEGLQALLRDQDAAPPGAAAFYAP GPLPPGRGSEHYSGDSDASSPRSNCSDGMMDYSGPPSGPRRQNGY DTAYYSEAVRESRPGKSAAVSSLDCLSSIVERISIDSPAAPALLL ADAPPESPPGPPEGASLSDTEQGTQTPSPDAAPQCPAGSNPNAIY QVL (GenBank Accession No. AAA39798.1, which is hereby incorporated by reference in its entirety) NF-KB MDELFPLIFPAEQPKQRGMRFRYKCEGRSAGSIPGERSTDTTKTH 52 (*Homo sapiens*) PTIKINGYTGPGTVRISLVTKDPPHRPHPHELVGKDCRDGFYEAE LCPDRCIHSFQNLGIQCVKKRDLEQAISQRIQTNNNPFQVPIEEQ RGDYDLNAVRLCFQVTVRDPSGRPLRLPPVLSHPIFDNRAPNTAE LKICRVNRNSGSCLGGDEIFLLCDKVQKEDIEVYFTGPGWEARGS FSQADVHRQVAIVFRTPPYADPSLQAPVRVSMQLRRPSDRELSEP MEFQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPP RRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQI

SQASALAPAPPQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPP
QAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDPAVF
TDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTAQR
PPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISS (GenBank Accession No. 2006293A, which is hereby incorporated by reference in its entirety) TetR MFISDKVSSMTKLQPNTVIRAALDLLNEVGVDGLTTRKLAERLGV 53 (*Proteobacteria*)
QQPALYWHFRNKRALLDALAEAMLAENHTHSVPRADDDWRSFLIG
NARSFRQALLAYRDGARIHAGTRPGAPQMETADAQLRFLCEAGFS
AGDAVNALMTISYFTVGAVLEEQAGDSDAGERGGTVEQAPLSPLL
RAAIDAFDEAGPDAAFEQGLAVIVDGLAKRRLVVRNVEGPRKGDD (GenBank Accession No. WP_000470728.1, which is hereby incorporated by reference in its entirety)

(40) Additional exemplary transcription factors are identified in Table 5 below.

(41) TABLE-US-00005 TABLE 5 Additional Exemplary Transcription Factors Transcription Factor Family Transcriptions Factors Basic Helix- AHR, ARNT/HIF-1 beta , ASCL1/Mash1, ASCL2/Mash2, CLOCK, Loop-Helix DEC2, HAND1, HAND2, HES-1, HES-4, HIF-1 alpha/HIF1A, HIF-2, (bHLH) Family alpha/EPAS1, c-Maf, Max, MESP1, MITF, MLX, Mxi1, c-Myc, MYCL1/L-Myc, MYF-5, MyoD, Myogenin, NeuroD1, NeuroD2, Neurogenin-1, Neurogenin-2, Neurogenin-3, Olig1, Olig2, Olig3, SCL/Tal1, SREBP2, TCF-12/HTF4, TFEB, Twist-1 Twist-2, UTF1 Basic Leucine ATF1, ATF2, ATF4, BACH1, BATF, BATF3, c-Fos, CEBP alpha, Zipper (bZIP) CEBP epsilon, CREB, FosB/G0S3, FRA-1, GADD153, HSF1, HSF2, Family HSF4, c-Jun, JunB, JunD, c-Maf, MafB, MafF, MafG, MafK, Max, MITF, MLX, Mxi1, MYB, c-Myc, MYCL1/L-Myc, NFIL3/E4BP4, Nrf1, Nrf2, NRL, OASIS/CREB3L1, SREBP2, TSC22, XBP1 ETS (E-twenty ELF3, Ets-1, ETV1, ETV2/ER71, ETV5, ETV6, FLI1, PU.1/Spi-1, six) Family Spi-B Forkhead Domain FoxC1, FoxC2, FoxD3, FoxF1, FoxF2, FoxH1, FoxJ1, FoxJ3, FoxK1, Family FOXL2, FoxM1, FoxN1, FoxO1/FKHR, FoxO3, FoxP1, FoxP2, FoxP3, FoxP4, HNF-3 alpha/FoxA1, HNF-3 beta/FoxA2 GATA Family GATA-1, GATA-2, GATA-3, GATA-4, GATA-5, GATA-6, TRPS1 Hypoxia HIF-1, HIF-2, HIF-3, ARNT/HIF-1 beta Inducible Factors (HIFs) Family High Mobility HMGA1B, HMGA2, HMGB1/HMG-1, HMGB3, HMGN1, LEF1, Group (HMG) SOX1, SOX2, SOX3, SOX5, SOX6, SOX7, SOX9, SOX10, SOX11, • Family SOX15, SOX17, SOX18, SOX21, TCF7/TCF1, TCF7L1/TCF3 Homeodomain ADNP, ARX, ATBF1/ZFHX3, CDX2, CDX4, CRX, DLX5, DUX4, (Hox) Family DUX4/DUX4c, DUX4c, EMX2, GBX2, Goosecoid, HHEX, HNF- 6/ONECUT1, HOXA1, HOXB1, HOXB7, HOXB13, HOXD10, Islet- 1, Islet-2, LHX5, LIM1, MSX1, MSX2, Nanog, NKX2.2, NKX2.5, NKX3.1, NKX6.1, Oct-1, Oct-3/4, Oct-4A, Oct-4B, ONECUT2/OC-2, Otx2, PDX-1/IPF1, PHOX2B, PITX2, POU3F2, Prox1, SATB1, TCF- 2/HNF-1 beta, TCF-3/E2A, TGIF1, TTF-1/NKX2-1, VSTM2L, ZEB1 Immunoglobulin- CSL, NFκB, p50 (NFκB1), p52 (NFκB2), p53, p63/TP73L, Like Domain NFκBp65/RelA, RelB, c-Rel, STAT (STAT1, STAT2, STAT3, Family STAT4, STAT5a/b, STAT5a, STAT5b, STAT6) Interferon- IRF1, IRF2, IRF3, IRF4, IRF5, IRF6, IRF8 Regulatory Factor (IRF) Family Kruppel-like KLF2, KLF4, KLF5, KLF6, KLF10, KLF12, KLF17 Family Paired Box (Pax) Pax2, Pax3, Pax4, Pax5/BSAP, Pax6, Pax7 Family Mothers against FOXL2, Smad1, Smad2, Smad2/3, Smad3, Smad4, Smad5, Smad7, decapentaplegic Smad8, Smad9 homolog (Smad) Family Additional AP-2 beta, AP-2 gamma, AP-2 epsilon, Autoimmune Regulator/AIRE, Transcription BLIMP1/PRDM1, C1D, DACH2, DC-SCRIPT/ZNF366, DIDO1, E2F- Factors 1, E2F-2, E2F-4, EGR1, GLI-1, GLI-2, GLI-3, HNF-4 alpha/NR2A1, HNF-4 gamma/NR2A2, LMO2, LMO4, LPP, MEF2C, PREB, RFX6, Teneurin-1, Teneurin-2, Teneurin-4, TFCP2L1, ZSCAN21

(42) RNA-regulated destabilization domains are amino acid sequences that, when functionally coupled to a protein of interest, modulate the stability of the protein of interest in a RNA-dependent manner. In some embodiments, when the RNA-regulated destabilization domain is fused to a

protein of interest, the RNA-regulated destabilization domain mediates protein degradation. In accordance with such embodiments, the protein destabilization function of the RNA-regulated destabilization domain is impeded when it binds to a specific RNA molecule (e.g., an aptamer).
(43) In some embodiments, the RNA-regulated destabilization domain comprises a bifunctional peptide comprising an RNA-binding domain and a degron peptide. The RNA-binding domain may be any peptide to which an RNA molecule can bind, where such binding sterically inhibits the interaction of the degron peptide with a proteosomal pathway component (e.g., an E3 ubiquitin ligase). Thus, in some embodiments, the RNA-binding domain is MDARTRRRERRAEKQAQWKAAN (lambdaN; SEQ ID NO: 123), which is derived from the lambda bacteriophage antiterminator protein N. In accordance with such embodiments, the RNA-binding domain is specific for BoxB (SEQ ID NO: 124): GGGCCCUGAAGAAGGGCCC (see, e.g., "NMR Structure of the Bacteriophage Lambda N Peptide/boxB RNA Complex: Recognition of a GNRA Fold by an Arginine-Rich Motif," *Cell* 93(2):289-299 (1998), which is hereby incorporated by reference in its entirety).
(44) In other embodiments, the RNA-binding domain is DTRQARRNRRRRWRERQRAAAAR (HIV-1 Rev; SEQ ID NO: 125), which is derived from HIV-1 Rev peptide. In accordance with such embodiments, the RNA-binding domain is specific for RRE RNA (SEQ ID NO: 126): GGUCUGGGCGCAGCGCAAGCUGCGGACAGGCC (see, e.g., Battiste et al., "Alpha Helix—RNA Major Groove Recognition in an HIV-1 Rev Peptide—RRE RNA Complex," *Science* 273:1547-1551 (1996), which is hereby incorporated by reference in its entirety).
(45) The RNA-regulated destabilization domain may comprise a bifunctional peptide comprising a lentiviral transactivator of transcription (Tat) peptide and a degron peptide.
(46) In some embodiments, the lentiviral Tat peptide is a bovine immunodeficiency virus Tat peptide. In other embodiments, the lentiviral Tat peptide is a human immunodeficiency virus Tat peptide.
(47) According to some embodiments, the Tat peptide has the sequence of RKKRRQRRR (SEQ ID NO: 129). See, e.g., Yamamoto et al., "A Novel RNA Motif that Binds Efficiently and Specifically to the Ttat Protein of HIV and Inhibits the Trans-Activation by Tat of Transcription In Vitro and In Vivo," *Genes Cells* 5:371-388 (2000), which is hereby incorporated by reference in its entirety.
(48) According to some embodiments, the Tat peptide has the consensus sequence of SEQ ID NO: 54 as follows: XXXXXXXXXXXXXXX, where X at position 1 can be S or A; X at position 2 can be G or A; X at position 3 can be P or A; X at position 4 can be R or K; X at position 5 can be P, A, I, Y, K, or R; X at position 6 can be R, K, V, or Y; X at position 7 can be G, A, or R; X at position 8 can be T or A; X at position 9 can be R or K; X at position 10 can be G or A; X at position 11 can be K or A; X at position 12 can be G or A; X at position 13 can be R or K; X at position 14 can be I or A; X at position 15 can be R, K, Y, or G; and X at position 16 can be R, K, V, T, or Y. See, e.g., Athanassiou et al., "Structural Mimicry of Retroviral Tat Proteins by Constrained β-Hairpin Peptidomimetics: Ligands with High Affinity and Selectivity for Viral TAR RNA Regulatory Elements," *J. Am. Chem. Soc.* 126:6906-6913 (2004); Chen & Frankel, "A Peptide Interaction in the Major Groove of RNA Resembles Protein Interactions in the Minor Groove of DNA," *Proc. Natl. Acad. Sci. USA* 92:5077-5081 (1995); and Koren et al., "The Eukaryotic Proteome is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons," *Cell* 173:1622-1635 (2018), which are hereby incorporated by reference in their entirety). For example, the Tat peptide may have the amino acid sequence of SEQ ID NO: 55 as follows: SGPRPRGTRGKGRIRR.
(49) In some embodiments, the lentiviral Tat peptide comprises an RNA binding site. The RNA binding site may correspond to amino acid residues 4-17 of SEQ ID NO: 54 or amino acid residues 4-17 of SEQ ID NO: 55.
(50) In some embodiments, the RNA binding site is specific for an RNA aptamer. An aptamers is a nucleic acid molecule that binds with high affinity and specificity to a target. Nucleic acid aptamers

may be single-stranded, partially single-stranded, partially double-stranded, or double-stranded nucleotide sequences. Aptamers include, without limitation, defined sequence segments and sequences comprising nucleotides (e.g., ribonucleotides, nucleotide analogs, modified nucleotides, and nucleotides comprising backbone modifications, branchpoints, and non-nucleotide residues, groups, or bridges). Nucleic acid aptamers include partially and fully single-stranded and double-stranded nucleotide molecules and sequences; synthetic RNA, DNA, and chimeric nucleotides; hybrids; duplexes; heteroduplexes; and any ribonucleotide, deoxyribonucleotide, or chimeric counterpart thereof and/or corresponding complementary sequence, promoter, or primer-annealing sequence needed to amplify, transcribe, or replicate all or part of the aptamer molecule or sequence.

(51) As described herein, the RNA binding site is specific for an RNA aptamer having the consensus sequence of SEQ ID NO: 56 as follows: NNNNNSHSYWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. Thus, in some embodiments, the RNA aptamer has the sequence of wild-type TAR RNA (SEQ ID NO: 57) as follows: GGCUCGUGUAGCUCAUUAGCUCCGAGCC.

(52) According to some embodiments, the RNA binding site is specific for an RNA aptamer having the consensus sequence of SEQ ID NO: 58 as follows: NNNNNSHCYSWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. Thus, in some embodiments, the RNA aptamer has the sequence of TAR Variant-1 (SEQ ID NO: 59) as follows: GGCUCGUCUGAGCUCAUUAGCUCCGAGCC.

(53) In other embodiments, the RNA binding site is specific for an RNA aptamer having the consensus sequence of SEQ ID NO: 60 as follows: NNNNNSITYSWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. Thus, in some embodiments, the RNA aptamer has the sequence of TAR Variant-2 (Pepper; SEQ ID NO: 61) as follows: GGCUCGUUGAGCUCAUUAGCUCCGAGCC.

(54) In further embodiments, the RNA binding site is specific for an RNA aptamer having the sequence of HIV TAR (SEQ ID NO: 128) as follows: ACGAAGCUUGAUCCCGUUUGCCGGUCGAUCGCUUCGA.

(55) As used herein, the term “degron” or “degradation signal” or “degron peptide” refers to an amino acid element within a protein that is sufficient for recognition and degradation by a proteolytic system. In some embodiments, the degron is a ubiquitin-pathway degron. In accordance with such embodiments, the degron comprises a region specific for E3 binding (see, e.g., Ravid & Hochstrasser, “Diversity of Degradation Signals in the Ubiquitin-Proteasome System,” *Nat. Rev. Mol. Cell Biol.* 9:679-689 (2008), which is hereby incorporated by reference in its entirety).

(56) The degron peptide may be selected from a monopeptide, a dipeptide, a tripeptide, a tetrapeptide, a pentapeptide, a hexapeptide, a heptapeptide, or an octapeptide. Exemplary degron peptides are well known in the art and are listed in Table 6 below.

(57) TABLE-US-00006 TABLE 6 Exemplary Degron Peptides Degron Peptide Amino Acid Sequences Monopeptide P, E Dipeptide RG, GG, EE, AP, RP, NP, DP, CP, EP, QP, GP, HP, IP, LP, KP, MP, FP, PP, SP, TP, WP, YP, VP, SA, SR, SN, SD, SC, SE, SQ, SG, SH, SI, SL, SK, SM, SF, SP, SS, ST, SW, SY, SV, AN, RN, NN, DN, CN, EN, QN, GN, HN, IN, LN, KN, MN, FN, PN, SN, TN, WN, YN, VN, AD, RD, ND, DD, CD, ED, QD, GD, HD, ID, LD, KD, MD, FD, PD, SD, TD, WD, YD, VD, CA, CR, CN, CD, CC, CE, CQ, CG, CH, CI, CL, CK, CM, CF, CP, CS, CT, CW, CY, CV, AE, RE, NE, DE, CE, EE, QE, GE, HE, IE, LE, KE, ME, FE, PE, SE, TE, WE, YE, VE

(58) In some embodiments, the degron peptide is SEQ ID NO: 130 as follows: RRRG. In accordance with such embodiments, the destabilization domain has the sequence of HIV Tat-RRRG (SEQ ID NO: 127) as follows: RKKRRQRRRG.

(59) In other embodiments, the degron peptide is selected from the group consisting of FKBP12,

dihydrofolate reductase, and derivates thereof. See, e.g., Rakhit et al., "Evaluation of FKBP and DHFR Based Destabilizing Domains in *Saccharomyces Cerevisiae," Bioorg. Med. Chem. Lett.* 21:4965-4968 (2011) and Iwamoto et al., "A General Chemical Method to Regulate Protein Stability in the Mammalian Central Nervous System," *Chem. Biol.* 17:981-988 (2010), which are hereby incorporated by reference in their entirety). In some embodiments, the FKBP12 is a human FKBP12. In some embodiments, the dihydrofolate reductase is an *E. coli* dehydrate reductase (ecDHFR). As described herein, aptamers that selectively bind to FKBP12, DHFR, or derivatives thereof may be used to confer stability to a protein of interest comprising FKBP12, ecDHFR, or a derivative thereof as a fusion partner.

(60) In some embodiments, the destabilization domain has the consensus sequence of SEQ ID NO: 62 as follows: XXXXXXXXXXXXXXXXx, where X at position 1 can be S or A; X at position 2 can be G or A; X at position 3 can be P or A; X at position 4 can be R or K; X at position 5 can be P, A, I, Y, K, or R; X at position 6 can be R, K, V, or Y; X at position 7 can be G, A, or R; X at position 8 can be T or A; X at position 9 can be R or K; X at position 10 can be G or A; X at position 11 can be K or A; X at position 12 can be G or A; X at position 13 can be R or K; X at position 14 can be I or A; X at position 15 can be R, K, Y, or G; X at position 16 can be R, K, V, T, or Y; X at position 17 can be any amino acid but preferably R, G, E, S, or C; and x at position 18 is optional and can be any amino acid, but preferably G, E, O, N, D, or E.

(61) In some embodiments the destabilization domain has the sequence of tDeg (SEQ ID NO: 63) as follows: SGPRPRGTRGKGRRIRRRG.

(62) The nucleic acid molecule described herein may further comprise a third nucleic acid sequence encoding a second protein of interest, wherein the third nucleic acid sequence is located between the first nucleic acid sequence and second nucleic acid sequence. Suitable proteins of interest are described in more detail above and include, without limitation, a fluorescent protein, a bioluminescent protein, an enzyme, or a transcriptional regulator.

(63) Another aspect of the invention relates to a nucleic acid molecule encoding a lentiviral transactivator of transcription (Tar) RNA aptamer sequence.

(64) In some embodiments, the lentiviral transactivator of transcription (Tar) RNA aptamer sequence is a bovine immunodeficiency virus (BIV) Tar sequence. In other embodiments, the lentiviral transactivator of transcription (Tar) RNA sequence is a human immunodeficiency virus (HIV) Tar sequence.

(65) According to some embodiments, the nucleic acid molecule encoding the lentiviral Tar RNA sequence is a DNA molecule according to the consensus sequence of SEQ ID NO: 64 as follows: NNNNNSHSYWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or T; S can be C or G; H can be A, C, or T; Y can be C or T; W can be A or T; B can be C, G, or T; M can be A or C; and D can be A, G, or T. For example, the nucleic acid molecule encoding the lentiviral Tar RNA sequence may be a DNA molecule encoding wild-type TAR RNA as follows: GGCTCGTGTAGCTCATTAGCTCCGAGCC (SEQ ID NO: 65).

(66) According to some embodiments, the nucleic acid molecule encoding the lentiviral TAR RNA sequence is a DNA molecule according to the consensus sequence of SEQ ID NO: 66 as follows: NNNNNSHCYSWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or T; S can be C or G; H can be A, C, or T; Y can be C or T; W can be A or T; B can be C, G, or T; M can be A or C; and D can be A, G, or T. For example, the nucleic acid molecule encoding the lentiviral Tar RNA sequence may be a DNA molecule encoding TAR Variant-1 as follows: GGCTCGTCTGAGCTCATTAGCTCCGAGCC (SEQ ID NO: 67).

(67) According to some embodiments, the nucleic acid molecule encoding the lentiviral TAR RNA sequence is a DNA molecule according to the consensus sequence of SEQ ID NO: 68 as follows: NNNNNSITYSWSBMNNNNDSBHBSNNNNN, where N can be A, C, G, or T; S can be C or G; H can be A, C, or T; Y can be C or T; W can be A or T; B can be C, G, or T; M can be A or C; and D can be A, G, or T. For example, the nucleic acid molecule encoding the lentiviral Tar RNA

sequence may be a DNA molecule encoding TAR Variant-2 (Pepper) as follows:
GGCTCGTTGAGCTCATTAGCTCCGAGCC (SEQ ID NO: 69).
(68) Suitable additional lentiviral transactivator of transcription (Tar) RNA aptamer sequences of the present application are shown in Table 7 below.
(69) TABLE-US-00007 TABLE 7 TAR RNA Sequences SEQ ID TAR RNA Sequence NO: (Pepper).sub.10tag
GGCUCGUCUGAGCUCAUUAGCUCCGAGCCGUCCAGCGCAAACUAU 70
UACGAAAAACAUCCGACGGGCUCGUUGAGCUCAUUAGCUCCGAGC
CCGCUGCGGAAAACCUCACAAAAACACGACAAACGGGCUCGUUGA
GCUCAUUAGCUCCGAGCCCGCCGACAACCCACAAACUUACAACCA
GGCAAACGGCUCGUCUGAGCUCAUUAGCUCCGAGCCGUAUCAAGA
CCGAACGGCGCAAGAUAUUGACACGGGCUCGUUGAGCUCAUUAGC
UCCGAGCCCGACCUCGCUAGAUAUGUUAGGUUCUUAGGCAUUGGC
UCGUUGAGCUCAUUAGCUCCGAGCCAAAGAUCGACUGCAAUUCCG
AUUAGACGUACACGGCUCGUCUGAGCUCAUUAGCUCCGAGCCGAU
CCAACCUACUUCCUCCAUAACUAACCUCCGGCUCGUUGAGCUCAU
UAGCUCCGAGCCGAUCAUAACGCAAUACCGUACACUGUCCAAUCC
GGCUCGUUGAGCUCAUUAGCUCCGAGCCGGACAACCAAUCGACAU
ACAUCACACCACAACUCGGCUCGUCUGAGCUCAUUAGCUCCGAGC C (F30-
UUGCCAUGUGUAUGUGGGAUGCGUUGCCACGUUUCCCACAUACUC 71
1xPepper).sub.10tag
UGAUGAUCCGCUAGCAAAGGCUCGUCUGAGCUCAUUAGCUCCGAG
CCCGAGGUACCGGAUCAUUCAUGGCAAGUCCAGCGCAAUCUAUUA
CGAAAAUCAUCCGACGUCGCGAUGUCUAUGCGGGAUGCGUUGCCA
CGUUUCCCGCAUAGUCUGAUCAUCCGCUAGCAAAGGCUCGUUGAG
CUCAUUAGCUCCGAGCCCGAGGUACCGGAUGAUUCAUCGCGACGC
UGCGGAAAAUCUCACAAAAUCACGUCAAACGUCGCCGUGUGUGUG
UAGGAUGCGUUGCCACGUUUCCUACACACUCUGACGAUCCGCUAG
CAAAGGCUCGUUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGAU
CGUUCACGGCGACGCCGAUAAUCCACAUACUUACAAUCAGGCAAU
CUUGCCAUGUGUAUGUGGGAUGCGUUGCCACGUUUCCCACAUACU
CUGAUGAUCCGCUAGCAAAGGCUCGUUGAGCUCAUUAGCUCCGAG
CCCGAGGUACCGGAUCAUUCAUGGCAAGUAUCAAGAUCGAACGGC
GCAAGAUAUUGUCACGUCGCGAUGUCUAUGCGGGAUGCGUUGCCA
CGUUUCCCGCAUAGUCUGAUCAUCCGCUAGCAAAGGCUCGUCUGA
GCUCAUUAGCUCCGAGCCCGAGGUACCGGAUGAUUCAUCGCGACG
UCCUCGCUAGAUAUGUUAGGUUCUUAGGCAUUUCGCCGUGUGUGU
GUAGGAUGCGUUGCCACGUUUCCUACACACUCUGACGAUCCGCUA
GCAAAGGCUCGUUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGA
UCGUUCACGGCGAAAAGAUCGUCUGCAAUUCCGAUUAGACGUACA
CUUGCCAUGUGUAUGUGGGAUGCGUUGCCACGUUUCCCACAUACU
CUGAUGAUCCGCUAGCAAAGGCUCGUUGAGCUCAUUAGCUCCGAG
CCCGAGGUACCGGAUCAUUCAUGGCAAGAUCCAAGCUACUUCCUC
CAUACCUAUCCUCCUCGCGAUGUCUAUGCGGGAUGCGUUGCCACG
UUUCCCGCAUAGUCUGAUCAUCCGCUAGCAAAGGCUCGUUGAGCU
CAUUAGCUCCGAGCCCGAGGUACCGGAUGAUUCAUCGCGAGAUCA
UAACGCAAUACCGUACACUGUCCAAUCCUCGCCGUGUGUGUGUAG
GAUGCGUUGCCACGUUUCCUACACACUCUGACGAUCCGCUAGCAA
AGGCUCGUCUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGAUCG
UUCACGGCGAGGAUAAUCAAUCCACAUACAUCACACCACAAUUCU

UGCCAUGUGUAUGUGGGAUGCGUUGCCACGUUUCCCACAUACUCU
GAUGAUCCGCUAGCAAAGGCUCGUCUGAGCUCAUUAGCUCCGAGC
CCGAGGUACCGGAUCAUUCAUGGCAA (Pepper).sub.20-tag
GGCUCGUCUGAGCUCAUUAGCUCCGAGCCGUCCAGCGCAAACUAU 72
UACGAAAAACAUCCGACGGGCUCGUUGAGCUCAUUAGCUCCGAGC
CCGCUGCGGAAAACCUCACAAAAACACGACAAACGGGCUCGUUGA
GCUCAUUAGCUCCGAGCCCGCCGACAACCCACAAACUUACAACCA
GGCAAACGGCUCGUCUGAGCUCAUUAGCUCCGAGCCGUAUCAAGA
CCGAACGGCGCAAGAUAUUGACACGGGCUCGUUGAGCUCAUUAGC
UCCGAGCCCGACCUCGCUAGAUAUGUUAGGUUCUUAGGCAUUGGC
UCGUUGAGCUCAUUAGCUCCGAGCCAAAGAUCGACUGCAAUUCCG
AUUAGACGUACACGGCUCGUCUGAGCUCAUUAGCUCCGAGCCGAU
CCAACCUACUUCCUCCAUAACUAACCUCCGGCUCGUUGAGCUCAU
UAGCUCCGAGCCGAUCAUAACGCAAUACCGUACACUGUCCAAUCC
GGCUCGUUGAGCUCAUUAGCUCCGAGCCGGACAACCAAUCGACAU
ACAUCACACCACAACUCGGCUCGUCUGAGCUCAUUAGCUCCGAGC
CGAAUUGGUCGUUCUUCUUGGCGGCCGCUCGACUAAGGUGACAAC
UGGACAAACCCUCGGCUCGUUGAGCUCAUUAGCUCCGAGCCGACU
CUCACCAACAAGACAAAAACUACUCUUCUAGGCUCGUUGAGCUCA
UUAGCUCCGAGCCUAAACACUCAAGCAUACAUUGUGCCUAUUUCU
UGGCUCGUCUGAGCUCAUUAGCUCCGAGCCAUGCUCUCACGAAUU
UCAAAACACGGACAAGGGGCUCGUUGAGCUCAUUAGCUCCGAGCC
CGUUCCACGUCCAAUACGAUUACUUACCUUUCGGGCUCGUUGAGC
UCAUUAGCUCCGAGCCCGCAGCUACAUCACUUCCACUCAGGACAU
UCAAGGGCUCGUCUGAGCUCAUUAGCUCCGAGCCCUCCACAAGUC
UCAACCACAGAAACUACCAAAUGGGCUCGUUGAGCUCAUUAGCUC
CGAGCCCACUCCUACCUCAAACCUCUUCCCACAAAACUGGGGCUC
GUUGAGCUCAUUAGCUCCGAGCCCCCAUUCCAACAUACCAAAUCA
AAAACAAUUACUGGCUCGUCUGAGCUCAUUAGCUCCGAGCCAGCC
CACAUCUCUCACUACUAUCAAAAACCAAACGGCUCGUUGAGCUCA
UUAGCUCCGAGCC (F30-
UUGCCAUGUGUAUGUGGGAAGCGUAGAAAGGCUCGUUGAGCUCAU 73
2xPepper).sub.10tag
UAGCUCCGAGCCCGACUACGUUUCCCACAUACUCUGAUGAUCCGC
UAGCAAAGGCUCGUCUGAGCUCAUUAGCUCCGAGCCCGAGGUACC
GGAUCAUUCAUGGCAAGUCCAGCGCAAUCUAUUACGAAAAUCAUC
CGACGUCGCGAUGUCUAUGCGGGAAGCGUAGAAAGGCUCGUCUGA
GCUCAUUAGCUCCGAGCCCGACUACGUUUCCCGCAUAGUCUGAUC
AUCCGCUAGCAAAGGCUCGUUGAGCUCAUUAGCUCCGAGCCCGAG
GUACCGGAUGAUUCAUCGCGACGCUGCGGAAAAUCUCACAAAAUC
ACGUCAAACGUCGCCGUGUGUGUGUAGGAAGCGUAGAAAGGCUCG
UCUGAGCUCAUUAGCUCCGAGCCCGACUACGUUUCCUACACACUC
UGACGAUCCGCUAGCAAAGGCUCGUUGAGCUCAUUAGCUCCGAGC
CCGAGGUACCGGAUCGUUCACGGCGACGCCGAUAAUCCACAUACU
UACAAUCAGGCAAUCUUGCCAUGUGUAUGUGGGAAGCGUAGAAAG
GCUCGUUGAGCUCAUUAGCUCCGAGCCCGACUACGUUUCCCACAU
ACUCUGAUGAUCCGCUAGCAAAGGCUCGUUGAGCUCAUUAGCUCC
GAGCCCGAGGUACCGGAUCAUUCAUGGCAAGUAUCAAGAUCGAAC
GGCGCAAGAUAUUGUCACGUCGCGAUGUCUAUGCGGGAAGCGUAG
AAAGGCUCGUUGAGCUCAUUAGCUCCGAGCCCGACUACGUUUCCC

GCAUAGUCUGAUCAUCCGCUAGCAAAGGCUCGUCUGAGCUCAUUA
GCUCCGAGCCCGAGGUACCGGAUGAUUCAUCGCGACGUCCUCGCU
AGAUAUGUUAGGUUCUUAGGCAUUUCGCCGUGUGUGUGUAGGAAG
CGUAGAAAGGCUCGUUGAGCUCAUUAGCUCCGAGCCCGACUACGU
UUCCUACACACUCUGACGAUCCGCUAGCAAAGGCUCGUCUGAGCU
CAUUAGCUCCGAGCCCGAGGUACCGGAUCGUUCACGGCGAAAAGA
UCGUCUGCAAUUCCGAUUAGACGUACACUUGCCAUGUGUAUGUGG
GAAGCGUAGAAAGGCUCGUCUGAGCUCAUUAGCUCCGAGCCCGAC
UACGUUUCCCACAUACUCUGAUGAUCCGCUAGCAAAGGCUCGUUG
AGCUCAUUAGCUCCGAGCCCGAGGUACCGGAUCAUUCAUGGCAAG
AUCCAAGCUACUUCCUCCAUACCUAUCCUCCUCGCGAUGUCUAUG
CGGGAAGCGUAGAAAGGCUCGUCUGAGCUCAUUAGCUCCGAGCCC
GACUACGUUUCCCGCAUAGUCUGAUCAUCCGCUAGCAAAGGCUCG
UUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGAUGAUUCAUCGC
GAGAUCAUAACGCAAUACCGUACACUGUCCAAUCCUCGCCGUGUG
UGUGUAGGAAGCGUAGAAAGGCUCGUCUGAGCUCAUUAGCUCCGA
GCCCGACUACGUUUCCUACACACUCUGACGAUCCGCUAGCAAAGG
CUCGUUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGAUCGUUCA
CGGCGAGGAUAAUCAAUCCACAUACAUCACACCACAAUUCUUGCC
AUGUGUAUGUGGGAAGCGUAGAAAGGCUCGUCUGAGCUCAUUAGC
UCCGAGCCCGACUACGUUUCCCACAUACUCUGAUGAUCCGCUAGC
AAAGGCUCGUCUGAGCUCAUUAGCUCCGAGCCCGAGGUACCGGAU
CAUUCAUGGCAA

(70) In some embodiments, the nucleic acid molecule further encodes at least one additional RNA aptamer. Thus, in some embodiments, the nucleic acid molecule may encode a lentiviral transactivator of transcription (Tar) RNA aptamer operably coupled to at least one additional RNA aptamer. The at least one additional aptamer may be a S-adenosylmethionine (SAM)-binding aptamer. For example, the nucleic acid molecule may encodes a SAM-binding aptamer operably linked to the lentiviral transactivator of transcription (Tar) RNA aptamer. As described herein, binding of SAM to its aptamer promotes folding of other linked aptamers, such as Pepper. In this way, the expressed RNA is a “sensor” which couples SAM levels to Pepper folding.

(71) Also contemplated are nucleic acid molecules encoding a protein-binding RNA sequence. Thus, in some embodiments, the nucleic acid molecule encodes a non-lentiviral transactivator of transcription (Tar) RNA sequence. In accordance with such embodiments, the protein-binding RNA sequence is BoxB or RRE.

(72) Some embodiments of the present application relate to a vector comprising a nucleic acid molecule described herein (i.e., a nucleic acid molecule encoding an RNA-regulated fusion protein and/or a lentiviral transactivator of transcription (Tar) RNA sequence). As used herein, the term vector means any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which is capable of transferring gene sequences between cells. Thus, the term includes cloning and expression vectors, as well as viral vectors. The heterologous nucleic acid molecule is inserted into the expression system or vector in proper sense (5′ to 3′) orientation and correct reading frame. The vector contains the necessary elements for the transcription and/or translation of the inserted protein and/or RNA coding sequences of the present application.

(73) In one embodiment, the vector is a plasmid. Numerous vectors suitable for use in the compositions of the present application are known to those of skill in the art, and many are commercially available. The following vectors are provided by way of example; for eukaryotic cells: pcDNA3.1(+), Tornado (Litke & Jaffrey, “Highly Efficient Expression of Circular RNA Aptamers in Cells Using Autocatalytic Transcripts,” *Nat. Biotechnol.* 37(6):667-675(2019), which

is hereby incorporated by reference in its entirety), pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, and pSVLSV40 (Pharmacia). However, any other vector may be used so long as it is compatible with the cell.

(74) In another embodiment, the vector is a viral vector. Suitable viral expression vectors include, but are not limited to, viral vectors based on vaccinia virus; poliovirus; adenovirus (see, e.g., PCT Patent Application Publication Nos. WO 94/12649 to Gregory et al., WO 93/03769 to Crystal et al., WO 93/19191 to Haddada et al., WO 94/28938 to Wilson et al., WO 95/11984 to Gregory, and WO 95/00655 to Graham, which are hereby incorporated by reference in their entirety); adeno-associated virus (see, e.g., Flannery et al., "Efficient Photoreceptor-Targeted Gene Expression In Vivo by Recombinant Adeno-Associated Virus," *PNAS* 94:6916-6921 (1997); Bennett et al., "Real-Time, Noninvasive In Vivo Assessment of Adeno-Associated Virus-Mediated Retinal Transduction," *Invest. Opthalmol. Vis. Sci.* 38:2857-2863 (1997); Jomary et al., "Nonviral Ocular Gene Transfer," *Gene Ther.* 4:683-690 (1997); Rolling et al., "Evaluation of Adeno-Associated Virus-Mediated Gene Transfer into the Rat Retina by Clinical Fluorescence Photography," *Hum. Gene. Ther.* 10:641-648 (1999); Ali et al., "Gene Transfer Into the Mouse Retina Mediated by an Adeno-Associated Viral Vector," *Hum. Mol. Genet.* 5:591-594 (1996); Samulski et al., "Helper-Free Stocks of Recombinant Adeno-Associated Viruses: Normal Integration Does not Require Viral Gene Expression," *J. Vir.* 63:3822-3828 (1989); Mendelson et al., "Expression and Rescue of a Nonselected Marker from an Integrated AAV Vector," *Virol.* 166:154-165 (1988); and Flotte et al., "Stable In Vivo Expression of the Cystic Fibrosis Transmembrane Conductance Regulator With an Adeno-Associated Virus Vector," *PNAS* 90:10613-10617 (1993), which are hereby incorporated by reference in their entirety); SV40; herpes simplex virus; human immunodeficiency virus (see, e.g., Miyoshi et al., "Stable and Efficient Gene Transfer into the Retina Using an HIV-Based Lentiviral Vector," *PNAS* 94:10319-10323 (1997), which is hereby incorporated by reference in its entirety); a retroviral vector, e.g., Murine Leukemia Virus, spleen necrosis virus, and vectors derived from retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, a lentivirus, human immunodeficiency virus, myeloproliferative sarcoma virus, and mammary tumor virus and the like.

(75) As described herein supra, the nucleic acid molecules encoding a protein of interest described herein may be inserted into a vector in the sense (i.e., 5′ to 3′) direction, such that the nucleic acid sequence encoding an RNA-regulated fusion protein is properly oriented for the expression of the encoded protein under the control of a promoter of choice. In some embodiments, the nucleic acid molecules encoding a RNA aptamer are inserted into the vector in the sense direction, such that the nucleic acid molecule encoding the RNA aptamer is properly oriented for the expression of a desired RNA aptamer. Single or multiple nucleic acid molecules may be ligated into an appropriate vector in this way, under the control of a suitable promoter, to prepare a nucleic acid construct. A promoter is a DNA sequence which contains the binding site for RNA polymerase and initiates transcription of a downstream nucleic acid sequence. In one embodiment, the vector comprises a promoter. Thus, in some embodiments, the vector comprises a nucleic acid molecule encoding a lentiviral transactivator of transcription (Tar) aptamer (e.g., Pepper) operably coupled to a promoter. In other embodiments, the vector comprises a nucleic acid molecule encoding a lentiviral transactivator of transcription (Tar) aptamer (e.g., Pepper) and at least one additional aptamer sequence (e.g., a S-adenosylmethionine (SAM)-binding aptamer) operably coupled to a promoter.

(76) The promoter may be a constitutively active promoter (i.e., a promoter that is constitutively in an active or "on" state), an inducible promoter (i.e., a promoter whose state, active or inactive state, is controlled by an external stimulus, e.g., the presence of a particular temperature, compound, or protein), a spatially restricted promoter (i.e., transcriptional control element, enhancer, etc.) (e.g., tissue specific promoter, cell type specific promoter, etc.), or a temporally restricted promoter (i.e., the promoter is in the "on" state or "off" state during specific stages of a biological process).

(77) Suitable promoters can be derived from viruses and can therefore be referred to as viral

promoters, or they can be derived from any organism, including prokaryotic or eukaryotic organisms. Suitable promoters can be used to drive expression by any RNA polymerase (e.g., RNA Polymerase I, RNA Polymerase II, RNA Polymerase III). The promoter may be a viral promoter. Exemplary promoters include, but are not limited to the SV40 early promoter, mouse mammary tumor virus long terminal repeat (LTR) promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a rous sarcoma virus (RSV) promoter, a human U6 small nuclear promoter (U6) (Miyagishi et al., "U6 Promoter-Driven siRNAs with Four Uridine 3′ Overhangs Efficiently Suppress Targeted Gene Expression in Mammalian Cells," *Nat. Biotechnol.* 20:497-500 (2002), which is hereby incorporated by reference in its entirety), an enhanced U6 promoter (e.g., Xia et al., "An Enhanced U6 Promoter for Synthesis of Short Hairpin RNA," *Nucleic Acids Res.* 31(17):e100 (2003), which is hereby incorporated by reference in its entirety), a human H1 promoter ("H1"), and the like. In some embodiments the promoter is a phage promoter, e.g., a T7 promoter that has been engineered to be expressed in a mammalian cell.

(78) Examples of inducible promoters include, but are not limited to T7 RNA polymerase promoter, T3 RNA polymerase promoter, isopropyl-beta-D-thiogalactopyranoside (IPTG)-regulated promoter, lactose induced promoter, heat shock promoter, tetracycline-regulated promoter, steroid-regulated promoter, metal-regulated promoter, estrogen receptor-regulated promoter, etc. Inducible promoters can therefore be regulated by molecules including, but not limited to, doxycycline, RNA polymerase, e.g., T7 RNA polymerase, an estrogen receptor, an estrogen receptor fusion, etc.

(79) In some embodiments, the promoter is a eukaryotic RNA polymerase promoter or a derivative thereof. Exemplary RNA polymerase II promoters include, without limitation, cytomegalovirus ("CMV"), phosphoglycerate kinase-1 ("PGK-1"), and elongation factor 1α ("EF1α") promoters. In yet another embodiment, the promoter is a eukaryotic RNA polymerase III promoter selected from the group consisting of U6, H1, 56, 7SK, and derivatives thereof.

(80) The RNA Polymerase promoter may be mammalian. Suitable mammalian promoters include, without limitation, human, murine, bovine, canine, feline, ovine, porcine, ursine, and simian promoters. In one embodiment, the RNA polymerase promoter sequence is a human promoter.

(81) According to one embodiment, the vector is a plasmid and has the sequence of pCMV-mCherry-(F30-2×Pepper).sub.10 (SEQ ID NO: 74; GenBank Accession No. MN052904.1, which is hereby incorporated by reference) as follows:

(82) TABLE-US-00008 1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG 151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG 201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG 351 CCCAACGACC CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT 451 AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA 751 ACAACTCCGC CCCCATTGAC GCAAATGGGC GGTAGGCGTG TACGGTGGGA 801 GGTCTATATA AGCAGAGCTC TCTGGCTAAC TAGAGAACCC ACTGCTTACT 851 GGCTTATCGA AATTAATACG ACTCACTATA GGGAGACCCA AGCTGGCTAG 901 CGTTTAAACT TAAGCTTGCC ACCATGGTGA

GCAAGGGCGA GGAGGATAAC 951 ATGGCCATCA TCAAGGAGTT CATGCGCTTC
AAGGTGCACA TGGAGGGCTC 1001 CGTGAACGGC CACGAGTTCG
AGATCGAGGG CGAGGGCGAG GGCCGCCCCT 1051 ACGAGGGCAC
CCAGACCGCC AAGCTGAAGG TGACCAAGGG TGGCCCCCTG 1101
CCCTTCGCCT GGGACATCCT GTCCCCTCAG TTCATGTACG GCTCCAAGGC 1151
CTACGTGAAG CACCCCGCCG ACATCCCCGA CTACTTGAAG CTGTCCTTCC 1201
CCGAGGGCTT CAAGTGGGAG CGCGTGATGA ACTTCGAGGA CGGCGGCGTG
1251 GTGACCGTGA CCCAGGACTC CTCCCTGCAG GACGGCGAGT
TCATCTACAA 1301 GGTGAAGCTG CGCGGCACCA ACTTCCCCTC
CGACGGCCCC GTAATGCAGA 1351 AGAAGACCAT GGGCTGGGAG
GCCTCCTCCG AGCGGATGTA CCCCGAGGAC 1401 GGCGCCCTGA
AGGGCGAGAT CAAGCAGAGG CTGAAGCTGA AGGACGGCGG 1451
CCACTACGAC GCTGAGGTCA AGACCACCTA CAAGGCCAAG AAGCCCGTGC
1501 AGCTGCCCGG CGCCTACAAC GTCAACATCA AGTTGGACAT
CACCTCCCAC 1551 AACGAGGACT ACACCATCGT GGAACAGTAC
GAACGCGCCG AGGGCCGCCA 1601 CTCCACCGGC GGCATGGACG
AGCTGTACAA GTAACTCGAG ATCCGTTACG 1651 GCCGGAATCA ATCGCTAATC
ACTCAACTTG CCATGTGTAT GTGGGAAGCG 1701 TAGAAAGGCT
CGTTGAGCTC ATTAGCTCCG AGCCCGACTA CGTTTCCCAC 1751 ATACTCTGAT
GATCCGCTAG CAAAGGCTCG TCTGAGCTCA TTAGCTCCGA 1801
GCCCGAGGTA CCGGATCATT CATGGCAAGT CCAGCGCAAT CTATTACGAA 1851
AATCATCCGA CGTCGCGATG TCTATGCGGG AAGCGTAGAA AGGCTCGTCT 1901
GAGCTCATTA GCTCCGAGCC CGACTACGTT TCCCGCATAG TCTGATCATC 1951
CGCTAGCAAA GGCTCGTTGA GCTCATTAGC TCCGAGCCCG AGGTACCGGA
2001 TGATTCATCG CGACGCTGCG GAAAATCTCA CAAAATCACG
TCAAACGTCG 2051 CCGTGTGTGT GTAGGAAGCG TAGAAAGGCT
CGTCTGAGCT CATTAGCTCC 2101 GAGCCCGACT ACGTTTCCTA CACACTCTGA
CGATCCGCTA GCAAAGGCTC 2151 GTTGAGCTCA TTAGCTCCGA
GCCCGAGGTA CCGGATCGTT CACGGCGACG 2201 CCGATAATCC ACATACTTAC
AATCAGGCAA TCTTGCCATG TGTATGTGGG 2251 AAGCGTAGAA
AGGCTCGTTG AGCTCATTAG CTCCGAGCCC GACTACGTTT 2301 CCCACATACT
CTGATGATCC GCTAGCAAAG GCTCGTTGAG CTCATTAGCT 2351
CCGAGCCCGA GGTACCGGAT CATTCATGGC AAGTATCAAG ATCGAACGGC
2401 GCAAGATATT GTCACGTCGC GATGTCTATG CGGGAAGCGT
AGAAAGGCTC 2451 GTTGAGCTCA TTAGCTCCGA GCCCGACTAC
GTTTCCCGCA TAGTCTGATC 2501 ATCCGCTAGC AAAGGCTCGT CTGAGCTCAT
TAGCTCCGAG CCCGAGGTAC 2551 CGGATGATTC ATCGCGACGT
CCTCGCTAGA TATGTTAGGT TCTTAGGCAT 2601 TTCGCCGTGT GTGTGTAGGA
AGCGTAGAAA GGCTCGTTGA GCTCATTAGC 2651 TCCGAGCCCG
ACTACGTTTC CTACACACTC TGACGATCCG CTAGCAAAGG 2701
CTCGTCTGAG CTCATTAGCT CCGAGCCCGA GGTACCGGAT CGTTCACGGC 2751
GAAAAGATCG TCTGCAATTC CGATTAGACG TACACTTGCC ATGTGTATGT 2801
GGGAAGCGTA GAAAGGCTCG TCTGAGCTCA TTAGCTCCGA GCCCGACTAC
2851 GTTTCCCACA TACTCTGATG ATCCGCTAGC AAAGGCTCGT TGAGCTCATT
2901 AGCTCCGAGC CCGAGGTACC GGATCATTCA TGGCAAGATC CAAGCTACTT
2951 CCTCCATACC TATCCTCCTC GCGATGTCTA TGCGGGAAGC GTAGAAAGGC
3001 TCGTCTGAGC TCATTAGCTC CGAGCCCGAC TACGTTTCCC GCATAGTCTG
3051 ATCATCCGCT AGCAAAGGCT CGTTGAGCTC ATTAGCTCCG AGCCCGAGGT
3101 ACCGGATGAT TCATCGCGAG ATCATAACGC AATACCGTAC ACTGTCCAAT
3151 CCTCGCCGTG TGTGTGTAGG AAGCGTAGAA AGGCTCGTCT GAGCTCATTA

3201 GCTCCGAGCC CGACTACGTT TCCTACACAC TCTGACGATC CGCTAGCAAA
3251 GGCTCGTTGA GCTCATTAGC TCCGAGCCCG AGGTACCGGA
TCGTTCACGG 3301 CGAGGATAAT CAATCCACAT ACATCACACC ACAATTCTTG
CCATGTGTAT 3351 GTGGGAAGCG TAGAAAGGCT CGTCTGAGCT
CATTAGCTCC GAGCCCGACT 3401 ACGTTTCCCA CATACTCTGA TGATCCGCTA
GCAAAGGCTC GTCTGAGCTC 3451 ATTAGCTCCG AGCCCGAGGT
ACCGGATCAT TCATGGCAAG AATTGGTCGT 3501 TCTTCTTGGC GGCCGCTCGA
CTAAATCACC GGTAATCTTC TTGTCCATCT 3551 AGACCTTATA AAGATCTTTG
TACAAGGGCC CGTTTAAACC CGCTGATCAG 3601 CCTCGACTGT GCCTTCTAGT
TGCCAGCCAT CTGTTGTTTG CCCCTCCCCC 3651 GTGCCTTCCT TGACCCTGGA
AAGGTGCCAC TCCCACTGTC CTTTCCTAAT 3701 AAAATGAGGA AATTGCATCG
CATTGTCTGA GTAGGTGTCA TTCTATTCTG 3751 GGGGGTGGGG
GTGGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT 3801
AGCAGGCATG CTGGGGATGC GGTGGGCTCT ATGGCTTCTG AGGCGGAAAG
3851 AACCAGCTGG GGCTCTAGGG GGTATCCCCA CGCGCCCTGT
AGCGGCGCAT 3901 TAAGCGCGGC GGGTGTGGTG GTTACGCGCA
GCGTGACCGC TACACTTGCC 3951 AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC
TTCCCTTCCT TTCTCGCCAC 4001 GTTCGCCGGC TTTCCCCGTC AAGCTCTAAA
TCGGGGGCTC CCTTTAGGGT 4051 TCCGATTTAG TGCTTTACGG CACCTCGACC
CCAAAAAACT TGATTAGGGT 4101 GATGGTTCAC GTAGTGGGCC
ATCGCCCTGA TAGACGGTTT TTCGCCCTTT 4151 GACGTTGGAG TCCACGTTCT
TTAATAGTGG ACTCTTGTTC CAAACTGGAA 4201 CAACACTCAA CCCTATCTCG
GTCTATTCTT TTGATTTATA AGGGATTTTG 4251 CCGATTTCGG CCTATTGGTT
AAAAAATGAG CTGATTTAAC AAAAATTTAA 4301 CGCGAATTAA
TTCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC 4351
AGGCTCCCCA GCAGGCAGAA GTATGCAAAG CATGCATCTC AATTAGTCAG
4401 CAACCAGGTG TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG
AAGTATGCAA 4451 AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
TAACTCCGCC 4501 CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCCG
CCCCATGGCT 4551 GACTAATTTT TTTTATTTAT GCAGAGGCCG AGGCCGCCTC
TGCCTCTGAG 4601 CTATTCCAGA AGTAGTGAGG AGGCTTTTTT
GGAGGCCTAG GCTTTTGCAA 4651 AAAGCTCCCG GGAGCTTGTA TATCCATTTT
CGGATCTGAT CAAGAGACAG 4701 GATGAGGATC GTTTCGCATG
ATTGAACAAG ATGGATTGCA CGCAGGTTCT 4751 CCGGCCGCTT
GGGTGGAGAG GCTATTCGGC TATGACTGGG CACAACAGAC 4801
AATCGGCTGC TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC
4851 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA
TGAACTGCAG 4901 GACGAGGCAG CGCGGCTATC GTGGCTGGCC
ACGACGGGCG TTCCTTGCGC 4951 AGCTGTGCTC GACGTTGTCA
CTGAAGCGGG AAGGGACTGG CTGCTATTGG 5001 GCGAAGTGCC
GGGGCAGGAT CTCCTGTCAT CTCACCTTGC TCCTGCCGAG 5051 AAAGTATCCA
TCATGGCTGA TGCAATGCGG CGGCTGCATA CGCTTGATCC 5101 GGCTACCTGC
CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC 5151
GTACTCGGAT GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG
5201 CATCAGGGGC TCGCGCCAGC CGAACTGTTC GCCAGGCTCA
AGGCGCGCAT 5251 GCCCGACGGC GAGGATCTCG TCGTGACCCA
TGGCGATGCC TGCTTGCCGA 5301 ATATCATGGT GGAAAATGGC CGCTTTTCTG
GATTCATCGA CTGTGGCCGG 5351 CTGGGTGTGG CGGACCGCTA
TCAGGACATA GCGTTGGCTA CCCGTGATAT 5401 TGCTGAAGAG
CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG 5451 GTATCGCCGC

TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC 5501 GAGTTCTTCT
GAGCGGGACT CTGGGGTTCG AAATGACCGA CCAAGCGACG 5551
CCCAACCTGC CATCACGAGA TTTCGATTCC ACCGCCGCCT TCTATGAAAG 5601
GTTGGGCTTC GGAATCGTTT TCCGGGACGC CGGCTGGATG ATCCTCCAGC 5651
GCGGGGATCT CATGCTGGAG TTCTTCGCCC ACCCCAACTT GTTTATTGCA 5701
GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA 5751
AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG 5801
TATCTTATCA TGTCTGTATA CCGTCGACCT CTAGCTAGAG CTTGGCGTAA 5851
TCATGGTCAT AGCTGTTTCC TGTGTGAAAT TGTTATCCGC TCACAATTCC 5901
ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG GGTGCCTAAT
5951 GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG
6001 TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC
AACGCGCGGG 6051 GAGAGGCGGT TTGCGTATTG GGCGCTCTTC
CGCTTCCTCG CTCACTGACT 6101 CGCTGCGCTC GGTCGTTCGG
CTGCGGCGAG CGGTATCAGC TCACTCAAAG 6151 GCGGTAATAC
GGTTATCCAC AGAATCAGGG GATAACGCAG GAAAGAACAT 6201
GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG GCCGCGTTGC
6251 TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA
6301 CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA
GATACCAGGC 6351 GTTTCCCCCT GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG
ACCCTGCCGC 6401 TTACCGGATA CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT
GGCGCTTTCT 6451 CATAGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG
TTCGCTCCAA 6501 GCTGGGCTGT GTGCACGAAC CCCCCGTTCA
GCCCGACCGC TGCGCCTTAT 6551 CCGGTAACTA TCGTCTTGAG
TCCAACCCGG TAAGACACGA CTTATCGCCA 6601 CTGGCAGCAG
CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG 6651
TGCTACAGAG TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGAA
6701 CAGTATTTGG TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA
6751 GTTGGTAGCT CTTGATCCGG CAAACAAACC ACCGCTGGTA GCGGTTTTTT
6801 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT
CAAGAAGATC 6851 CTTTGATCTT TTCTACGGGG TCTGACGCTC
AGTGGAACGA AAACTCACGT 6901 TAAGGGATTT TGGTCATGAG
ATTATCAAAA AGGATCTTCA CCTAGATCCT 6951 TTTAAATTAA AAATGAAGTT
TTAAATCAAT CTAAAGTATA TATGAGTAAA 7001 CTTGGTCTGA CAGTTACCAA
TGCTTAATCA GTGAGGCACC TATCTCAGCG 7051 ATCTGTCTAT TTCGTTCATC
CATAGTTGCC TGACTCCCCG TCGTGTAGAT 7101 AACTACGATA CGGGAGGGCT
TACCATCTGG CCCCAGTGCT GCAATGATAC 7151 CGCGAGACCC
ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA 7201
GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT
7251 CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA
7301 ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
7351 TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG
7401 AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC
7451 CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT
7501 ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT
7551 TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC
7601 GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
TACCGCGCCA 7651 CATAGCAGAA CTTTAAAAGT GCTCATCATT
GGAAAACGTT CTTCGGGGCG 7701 AAAACTCTCA AGGATCTTAC
CGCTGTTGAG ATCCAGTTCG ATGTAACCCA 7751 CTCGTGCACC CAACTGATCT

TCAGCATCTT TTACTTTCAC CAGCGTTTCT 7801 GGGTGAGCAA
AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC 7851
GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA 7901
GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT 7951
TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
8001 ACCTGACGTC

(83) According to one embodiment, the vector is a plasmid and has the sequence of pminiCMV-(mNeonGreen).sub.4-tDeg (SEQ ID NO: 75; GenBank Accession No. MN052905.1, which is hereby incorporated by reference) as follows:

(84) TABLE-US-00009 1 GACGGATCGG GAGATCTCCC GATCCCCTAT
GGTGCACTCT CAGTACAATC 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT
CTGCTCCCTG CTTGTGTGTT 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT
TTAAGCTACA ACAAGGCAAG 151 GCTTGACCGA CAATTGCATG AAGAATCTGC
TTAGGGTTAG GCGTTTTGCG 201 CTGCTTCGCG ATGTACGGGC CAGATATACG
CGTTGGTAGG CGTGTACGGT 251 GGGAGGCCTA TATAAGCAGA GCTAAGCTTG
CCACCATGGT GAGCAAGGGC 301 GAGGAGGATA ACATGGCCTC
TCTCCCAGCG ACACATGAGT TACACATCTT 351 TGGCTCCATC AACGGTGTGG
ACTTTGACAT GGTGGGTCAG GGCACCGGCA 401 ATCCAAATGA TGGTTATGAG
GAGTTAAACC TGAAGTCCAC CAAGGGTGAC 451 CTCCAGTTCT CCCCCTGGAT
TCTGGTCCCT CATATCGGGT ATGGCTTCCA 501 TCAGTACCTG CCCTACCCTG
ACGGGATGTC GCCTTTCCAG GCCGCCATGG 551 TAGATGGCTC CGGATACCAA
GTCCATCGCA CAATGCAGTT TGAAGATGGT 601 GCCTCCCTTA CTGTTAACTA
CCGCTACACC TACGAGGGAA GCCACATCAA 651 AGGAGAGGCC
CAGGTGAAGG GGACTGGTTT CCCTGCTGAC GGTCCTGTGA 701
TGACCAACTC GCTGACCGCT GCGGACTGGT GCAGGTCGAA GAAGACTTAC
751 CCCAACGACA AAACCATCAT CAGTACCTTT AAGTGGAGTT ACACCACTGG
801 AAATGGCAAG CGCTACCGGA GCACTGCGCG GACCACCTAC ACCTTTGCCA
851 AGCCAATGGC GGCTAACTAT CTGAAGAACC AGCCGATGTA CGTGTTCCGT
901 AAGACGGAGC TCAAGCACTC CAAGACCGAG CTCAACTTCA
AGGAGTGGCA 951 AAAGGCCTTT ACCGATGTGA TGGGCATGGA
CGAGCTGTAC AAGGGTGGAC 1001 ATATGGGCAC AGGGTCCACA
GGCGGTACCG GCGGAGTTTC CAAAGGAGAA 1051 GAAGACAATA
TGGCATCACT CCCCGCAACC CACGAGTTGC ATATTTTCGG 1101 TTCAATTAAT
GGAGTAGATT TCGATATGGT TGGCCAGGGA ACAGGAAACC 1151
CAAACGACGG ATATGAAGAG CTTAATCTCA AAAGTACCAA AGGCGATCTG
1201 CAATTTTCTC CGTGGATACT CGTGCCACAC ATTGGATACG GATTTCACCA
1251 ATATCTCCCG TATCCGGATG GAATGTCCCC CTTTCAAGCA GCAATGGTGG
1301 ACGGGAGTGG TTATCAGGTA CACAGAACCA TGCAGTTCGA
GGACGGGGCT 1351 TCTCTGACCG TAAATTATAG GTATACTTAT GAAGGCTCAC
ATATTAAGGG 1401 CGAAGCACAG GTTAAAGGAA CCGGGTTTCC
TGCGGATGGC CCCGTCATGA 1451 CTAATTCTCT GACAGCCGCA GATTGGTGTC
GCTCCAAAAA GACATACCCG 1501 AATGATAAGA CTATAATCTC AACATTCAAA
TGGTCCTATA CGACAGGCAA 1551 CGGGAAACGA TATAGATCCA
CGGCTCGAAC AACTTACACA TTCGCTAAAC 1601 CTATGGCCGC CAATTACCTC
AAAAATCAGC CCATGTATGT GTTTAGGAAA 1651 ACCGAATTGA AGCATTCTAA
AACGGAACTT AATTTTAAGG AATGGCAGAA 1701 GGCTTTCACA
GACGTAATGG GGATGGATGA ACTCTATAAA TCAGGTCTCG 1751
AGTCCTCAGG GGGAACGGGT GGGTCCGGAG GAGTTAGTAA AGGTGAAGAG
1801 GACAATATGG CAAGTTTGCC TGCGACTCAC GAGCTTCATA TCTTTGGGTC
1851 TATAAATGGC GTTGACTTCG ATATGGTTGG CCAAGGTACT GGCAACCCCA

1901 ATGACGGTTA CGAGGAGTTG AATCTCAAGT CCACAAAAGG
TGATCTTCAG 1951 TTCAGCCCTT GGATTCTCGT ACCTCATATT GGATATGGCT
TTCACCAGTA 2001 CCTTCCATAC CCAGACGGTA TGTCACCCTT TCAAGCTGCG
ATGGTGGATG 2051 GTTCCGGCTA TCAGGTCCAC CGAACGATGC
AATTCGAGGA CGGGGCCAGC 2101 CTCACCGTTA ATTATAGGTA CACCTATGAG
GGAAGTCACA TAAAGGGAGA 2151 AGCCCAAGTG AAAGGAACAG
GATTCCCAGC TGATGGTCCA GTAATGACGA 2201 ACTCCTTGAC
AGCGGCTGAC TGGTGTAGAA GCAAAAAGAC GTATCCTAAT 2251
GACAAGACCA TCATTAGCAC TTTCAAATGG AGTTATACCA CAGGAAACGG
2301 CAAACGGTAC AGAAGCACTG CTAGAACTAC CTACACTTTC
GCAAAGCCGA 2351 TGGCTGCAAA CTATTTGAAG AATCAGCCCA
TGTACGTGTT TCGAAAAACG 2401 GAACTTAAGC ACAGTAAGAC
TGAACTTAAT TTCAAGGAGT GGCAGAAGGC 2451 GTTCACGGAT
GTCATGGGTA TGGATGAACT GTATAAGGGA GGGTCTGGCA 2501
CTGGGGGCAC TGCCAGCAGC GGATCCGGTG GCGGTGTGAG CAAGGGCGAG
2551 GAGGATAACA TGGCCTCTCT CCCAGCGACA CATGAGTTAC ACATCTTTGG
2601 CTCCATCAAC GGTGTGGACT TTGACATGGT GGGTCAGGGC
ACCGGCAATC 2651 CAAATGATGG TTATGAGGAG TTAAACCTGA
AGTCCACCAA GGGTGACCTC 2701 CAGTTCTCCC CCTGGATTCT GGTCCCTCAT
ATCGGGTATG GCTTCCATCA 2751 GTACCTGCCC TACCCTGACG GGATGTCGCC
TTTCCAGGCC GCCATGGTAG 2801 ATGGCTCCGG ATACCAAGTC
CATCGCACAA TGCAGTTTGA AGATGGTGCC 2851 TCCCTTACTG TTAACTACCG
CTACACCTAC GAGGGAAGCC ACATCAAAGG 2901 AGAGGCCCAG
GTGAAGGGGA CTGGTTTCCC TGCTGACGGT CCTGTGATGA 2951
CCAACTCGCT GACCGCTGCG GACTGGTGCA GGTCGAAGAA GACTTACCCC
3001 AACGACAAAA CCATCATCAG TACCTTTAAG TGGAGTTACA
CCACTGGAAA 3051 TGGCAAGCGC TACCGGAGCA CTGCGCGGAC
CACCTACACC TTTGCCAAGC 3101 CAATGGCGGC TAACTATCTG
AAGAACCAGC CGATGTACGT GTTCCGTAAG 3151 ACGGAGCTCA
AGCACTCCAA GACCGAGCTC AACTTCAAGG AGTGGCAAAA 3201
GGCCTTTACC GATGTGATGG GCATGGACGA GCTGTACAAG GGCGGAAGAT
3251 CCGGTGGTGG TTCTGGTCCT CGTCCCCGTG GTACTCGTGG TAAAGGTCGC
3301 CGTATTCGTC GCCGCGGTTA ATCTAGAGGG CCCGTTTAAA CCCGCTGATC
3351 AGCCTCGACT GTGCCTTCTA GTTGCCAGCC ATCTGTTGTT TGCCCCTCCC
3401 CCGTGCCTTC CTTGACCCTG GAAAGGTGCC ACTCCCACTG TCCTTTCCTA
3451 ATAAAATGAG GAAATTGCAT CGCATTGTCT GAGTAGGTGT CATTCTATTC
3501 TGGGGGGTGG GGGTGGGGGC AGGACAGCAA GGGGGAGGAT
TGGGAAGACA 3551 ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT
CTATGGCTTC TGAGGCGGAA 3601 AGAACCAGCT GGGGCTCTAG
GGGGTATCCC CACGCGCCCT GTAGCGGCGC 3651 ATTAAGCGCG
GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG 3701
CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC 3751
ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG 3801
GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAAAAAA CTTGATTAGG 3851
GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT 3901
TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG 3951
AACAACACTC AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT 4001
TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA ACAAAAATTT 4051
AACGCGAATT AATTCTGTGG AATGTGTGTC AGTTAGGGTG TGGAAAGTCC
4101 CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC

4151 AGCAACCAGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC
AGAAGTATGC 4201 AAAGCATGCA TCTCAATTAG TCAGCAACCA
TAGTCCCGCC CCTAACTCCG 4251 CCCATCCCGC CCCTAACTCC GCCCAGTTCC
GCCCATTCTC CGCCCCATGG 4301 CTGACTAATT TTTTTTATTT ATGCAGAGGC
CGAGGCCGCC TCTGCCTCTG 4351 AGCTATTCCA GAAGTAGTGA
GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC 4401 AAAAAGCTCC
CGGGAGCTTG TATATCCATT TTCGGATCTG ATCAAGAGAC 4451 AGGATGAGGA
TCGTTTCGCA TGATTGAACA AGATGGATTG CACGCAGGTT 4501
CTCCGGCCGC TTGGGTGGAG AGGCTATTCG GCTATGACTG GGCACAACAG
4551 ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG
CGCAGGGGCG 4601 CCCGGTTCTT TTTGTCAAGA CCGACCTGTC
CGGTGCCCTG AATGAACTGC 4651 AGGACGAGGC AGCGCGGCTA
TCGTGGCTGG CCACGACGGG CGTTCCTTGC 4701 GCAGCTGTGC
TCGACGTTGT CACTGAAGCG GGAAGGGACT GGCTGCTATT 4751
GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC ATCTCACCTT GCTCCTGCCG
4801 AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA TACGCTTGAT
4851 CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA
TCGAGCGAGC 4901 ACGTACTCGG ATGGAAGCCG GTCTTGTCGA
TCAGGATGAT CTGGACGAAG 4951 AGCATCAGGG GCTCGCGCCA
GCCGAACTGT TCGCCAGGCT CAAGGCGCGC 5001 ATGCCCGACG
GCGAGGATCT CGTCGTGACC CATGGCGATG CCTGCTTGCC 5051 GAATATCATG
GTGGAAAATG GCCGCTTTTC TGGATTCATC GACTGTGGCC 5101
GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC TACCCGTGAT
5151 ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA
5201 CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG
5251 ACGAGTTCTT CTGAGCGGGA CTCTGGGGTT CGAAATGACC
GACCAAGCGA 5301 CGCCCAACCT GCCATCACGA GATTTCGATT
CCACCGCCGC CTTCTATGAA 5351 AGGTTGGGCT TCGGAATCGT TTTCCGGGAC
GCCGGCTGGA TGATCCTCCA 5401 GCGCGGGGAT CTCATGCTGG
AGTTCTTCGC CCACCCCAAC TTGTTTATTG 5451 CAGCTTATAA TGGTTACAAA
TAAAGCAATA GCATCACAAA TTTCACAAAT 5501 AAAGCATTTT TTTCACTGCA
TTCTAGTTGT GGTTTGTCCA AACTCATCAA 5551 TGTATCTTAT CATGTCTGTA
TACCGTCGAC CTCTAGCTAG AGCTTGGCGT 5601 AATCATGGTC ATAGCTGTTT
CCTGTGTGAA ATTGTTATCC GCTCACAATT 5651 CCACACAACA TACGAGCCGG
AAGCATAAAG TGTAAAGCCT GGGGTGCCTA 5701 ATGAGTGAGC
TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC 5751 AGTCGGGAAA
CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGCG 5801
GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA 5851
CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA
5901 AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC
AGGAAAGAAC 5951 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG
AACCGTAAAA AGGCCGCGTT 6001 GCTGGCGTTT TTCCATAGGC
TCCGCCCCCC TGACGAGCAT CACAAAAATC 6051 GACGCTCAAG
TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG 6101
GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC 6151
GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT 6201
CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC 6251
AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT
6301 ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
6351 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG

GTATGTAGGC 6401 GGTGCTACAG AGTTCTTGAA GTGGTGGCCT
AACTACGGCT ACACTAGAAG 6451 AACAGTATTT GGTATCTGCG
CTCTGCTGAA GCCAGTTACC TTCGGAAAAA 6501 GAGTTGGTAG
CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTTTT 6551 TTTGTTTGCA
AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA 6601
TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC 6651
GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC 6701
CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA 6751
AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG 6801
CGATCTGTCT ATTTCGTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG 6851
ATAACTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT 6901
ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC
6951 CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT
ATCCGCCTCC 7001 ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA
GTTCGCCAGT 7051 TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC
GTGGTGTCAC 7101 GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA
ACGATCAAGG 7151 CGAGTTACAT GATCCCCCAT GTTGTGCAAA
AAAGCGGTTA GCTCCTTCGG 7201 TCCTCCGATC GTTGTCAGAA
GTAAGTTGGC CGCAGTGTTA TCACTCATGG 7251 TTATGGCAGC ACTGCATAAT
TCTCTTACTG TCATGCCATC CGTAAGATGC 7301 TTTTCTGTGA CTGGTGAGTA
CTCAACCAAG TCATTCTGAG AATAGTGTAT 7351 GCGGCGACCG AGTTGCTCTT
GCCCGGCGTC AATACGGGAT AATACCGCGC 7401 CACATAGCAG
AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG 7451
CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC 7501
CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT 7551
CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG
7601 GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG
7651 AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA
7701 TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC
CCGAAAAGTG 7751 CCACCTGACG TC

(85) According to one embodiment, the vector is a plasmid and has the sequence of pCMV-CytERM-mCherry-(F30-2×Pepper).sub.10 (SEQ ID NO: 76; GenBank Accession No. MN052906.1, which is hereby incorporated by reference) as follows:

(86) TABLE-US-00010 1 GACGGATCGG GAGATCTCCC GATCCCCTAT
GGTGCACTCT CAGTACAATC 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT
CTGCTCCCTG CTTGTGTGTT 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT
TTAAGCTACA ACAAGGCAAG 151 GCTTGACCGA CAATTGCATG AAGAATCTGC
TTAGGGTTAG GCGTTTTGCG 201 CTGCTTCGCG ATGTACGGGC CAGATATACG
CGTTGACATT GATTATTGAC 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC
ATTAGTTCAT AGCCCATATA 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA
ATGGCCCGCC TGGCTGACCG 351 CCCAACGACC CCCGCCCATT GACGTCAATA
ATGACGTATG TTCCCATAGT 401 AACGCCAATA GGGACTTTCC ATTGACGTCA
ATGGGTGGAG TATTTACGGT 451 AAACTGCCCA CTTGGCAGTA CATCAAGTGT
ATCATATGCC AAGTACGCCC 501 CCTATTGACG TCAATGACGG TAAATGGCCC
GCCTGGCATT ATGCCCAGTA 551 CATGACCTTA TGGGACTTTC CTACTTGGCA
GTACATCTAC GTATTAGTCA 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC
AGTACATCAA TGGGCGTGGA 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT
CTCCACCCCA TTGACGTCAA 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG
GGACTTTCCA AAATGTCGTA 751 ACAACTCCGC CCCCATTGAC GCAAATGGGC
GGTAGGCGTG TACGGTGGGA 801 GGTCTATATA AGCAGAGCTC TCTGGCTAAC

TAGAGAACCC ACTGCTTACT 851 GGCTTATCGA AATTAATACG ACTCACTATA
GGGAGACCCA AGCTGGCTAG 901 CGTTTAAACT TGCCACCATG
GACCCTGTGG TGGTGCTGGG GCTCTGTCTC 951 TCCTGTTTGC TTCTCCTTTC
ACTCTGGAAA CAGAGCTATG GGGGAGGGAA 1001 ACTGGGCGGA
AGCGGAGGGA CGGGGGGTTC AGGAACTTCA GGGGGTGTGA 1051
GCAAGGGCGA GGAGGATAAC ATGGCCATCA TCAAGGAGTT CATGCGCTTC
1101 AAGGTGCACA TGGAGGGCTC CGTGAACGGC CACGAGTTCG
AGATCGAGGG 1151 CGAGGGCGAG GGCCGCCCCT ACGAGGGCAC
CCAGACCGCC AAGCTGAAGG 1201 TGACCAAGGG TGGCCCCCTG
CCCTTCGCCT GGGACATCCT GTCCCCTCAG 1251 TTCATGTACG GCTCCAAGGC
CTACGTGAAG CACCCCGCCG ACATCCCCGA 1301 CTACTTGAAG
CTGTCCTTCC CCGAGGGCTT CAAGTGGGAG CGCGTGATGA 1351
ACTTCGAGGA CGGCGGCGTG GTGACCGTGA CCCAGGACTC CTCCCTGCAG
1401 GACGGCGAGT TCATCTACAA GGTGAAGCTG CGCGGCACCA
ACTTCCCCTC 1451 CGACGGCCCC GTAATGCAGA AGAAGACCAT
GGGCTGGGAG GCCTCCTCCG 1501 AGCGGATGTA CCCCGAGGAC
GGCGCCCTGA AGGGCGAGAT CAAGCAGAGG 1551 CTGAAGCTGA
AGGACGGCGG CCACTACGAC GCTGAGGTCA AGACCACCTA 1601
CAAGGCCAAG AAGCCCGTGC AGCTGCCCGG CGCCTACAAC GTCAACATCA
1651 AGTTGGACAT CACCTCCCAC AACGAGGACT ACACCATCGT
GGAACAGTAC 1701 GAACGCGCCG AGGGCCGCCA CTCCACCGGC
GGCATGGACG AGCTGTACAA 1751 GTAACTCGAG ATCCGTTACG
GCCGGAATCA ATCGCTAATC ACTCAACTTG 1801 CCATGTGTAT GTGGGAAGCG
TAGAAAGGCT CGTTGAGCTC ATTAGCTCCG 1851 AGCCCGACTA
CGTTTCCCAC ATACTCTGAT GATCCGCTAG CAAAGGCTCG 1901 TCTGAGCTCA
TTAGCTCCGA GCCCGAGGTA CCGGATCATT CATGGCAAGT 1951
CCAGCGCAAT CTATTACGAA AATCATCCGA CGTCGCGATG TCTATGCGGG 2001
AAGCGTAGAA AGGCTCGTCT GAGCTCATTA GCTCCGAGCC CGACTACGTT
2051 TCCCGCATAG TCTGATCATC CGCTAGCAAA GGCTCGTTGA GCTCATTAGC
2101 TCCGAGCCCG AGGTACCGGA TGATTCATCG CGACGCTGCG
GAAAATCTCA 2151 CAAAATCACG TCAAACGTCG CCGTGTGTGT
GTAGGAAGCG TAGAAAGGCT 2201 CGTCTGAGCT CATTAGCTCC
GAGCCCGACT ACGTTTCCTA CACACTCTGA 2251 CGATCCGCTA
GCAAAGGCTC GTTGAGCTCA TTAGCTCCGA GCCCGAGGTA 2301
CCGGATCGTT CACGGCGACG CCGATAATCC ACATACTTAC AATCAGGCAA 2351
TCTTGCCATG TGTATGTGGG AAGCGTAGAA AGGCTCGTTG AGCTCATTAG 2401
CTCCGAGCCC GACTACGTTT CCCACATACT CTGATGATCC GCTAGCAAAG 2451
GCTCGTTGAG CTCATTAGCT CCGAGCCCGA GGTACCGGAT CATTCATGGC 2501
AAGTATCAAG ATCGAACGGC GCAAGATATT GTCACGTCGC GATGTCTATG 2551
CGGGAAGCGT AGAAAGGCTC GTTGAGCTCA TTAGCTCCGA GCCCGACTAC
2601 GTTTCCCGCA TAGTCTGATC ATCCGCTAGC AAAGGCTCGT CTGAGCTCAT
2651 TAGCTCCGAG CCCGAGGTAC CGGATGATTC ATCGCGACGT CCTCGCTAGA
2701 TATGTTAGGT TCTTAGGCAT TTCGCCGTGT GTGTGTAGGA AGCGTAGAAA
2751 GGCTCGTTGA GCTCATTAGC TCCGAGCCCG ACTACGTTTC CTACACACTC
2801 TGACGATCCG CTAGCAAAGG CTCGTCTGAG CTCATTAGCT CCGAGCCCGA
2851 GGTACCGGAT CGTTCACGGC GAAAAGATCG TCTGCAATTC CGATTAGACG
2901 TACACTTGCC ATGTGTATGT GGGAAGCGTA GAAAGGCTCG TCTGAGCTCA
2951 TTAGCTCCGA GCCCGACTAC GTTTCCCACA TACTCTGATG ATCCGCTAGC
3001 AAAGGCTCGT TGAGCTCATT AGCTCCGAGC CCGAGGTACC GGATCATTCA
3051 TGGCAAGATC CAAGCTACTT CCTCCATACC TATCCTCCTC GCGATGTCTA

3101 TGCGGGAAGC GTAGAAAGGC TCGTCTGAGC TCATTAGCTC
CGAGCCCGAC 3151 TACGTTTCCC GCATAGTCTG ATCATCCGCT
AGCAAAGGCT CGTTGAGCTC 3201 ATTAGCTCCG AGCCCGAGGT
ACCGGATGAT TCATCGCGAG ATCATAACGC 3251 AATACCGTAC ACTGTCCAAT
CCTCGCCGTG TGTGTGTAGG AAGCGTAGAA 3301 AGGCTCGTCT
GAGCTCATTA GCTCCGAGCC CGACTACGTT TCCTACACAC 3351 TCTGACGATC
CGCTAGCAAA GGCTCGTTGA GCTCATTAGC TCCGAGCCCG 3401
AGGTACCGGA TCGTTCACGG CGAGGATAAT CAATCCACAT ACATCACACC 3451
ACAATTCTTG CCATGTGTAT GTGGGAAGCG TAGAAAGGCT CGTCTGAGCT 3501
CATTAGCTCC GAGCCCGACT ACGTTTCCCA CATACTCTGA TGATCCGCTA 3551
GCAAAGGCTC GTCTGAGCTC ATTAGCTCCG AGCCCGAGGT ACCGGATCAT
3601 TCATGGCAAG AATTGGTCGT TCTTCTTGGC GGCCGCTCGA CTAAATCACC
3651 GGTAATCTTC TTGTCCATCT AGACCTTATA AAGATCTTTG TACAAGGGCC
3701 CGTTTAAACC CGCTGATCAG CCTCGACTGT GCCTTCTAGT TGCCAGCCAT
3751 CTGTTGTTTG CCCCTCCCCC GTGCCTTCCT TGACCCTGGA AAGGTGCCAC
3801 TCCCACTGTC CTTTCCTAAT AAAATGAGGA AATTGCATCG CATTGTCTGA
3851 GTAGGTGTCA TTCTATTCTG GGGGGTGGGG GTGGGGGCAG
GACAGCAAGG 3901 GGGAGGATTG GGAAGACAAT AGCAGGCATG
CTGGGGATGC GGTGGGCTCT 3951 ATGGCTTCTG AGGCGGAAAG
AACCAGCTGG GGCTCTAGGG GGTATCCCCA 4001 CGCGCCCTGT
AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA 4051
GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC 4101
TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC TTTCCCCGTC AAGCTCTAAA 4151
TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG CACCTCGACC 4201
CCAAAAAACT TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA 4251
TAGACGGTTT TTCGCCCTTT GACGTTGGAG TCCACGTTCT TTAATAGTGG 4301
ACTCTTGTTC CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT 4351
TTGATTTATA AGGGATTTTG CCGATTTCGG CCTATTGGTT AAAAAATGAG 4401
CTGATTTAAC AAAAATTTAA CGCGAATTAA TTCTGTGGAA TGTGTGTCAG 4451
TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA GTATGCAAAG
4501 CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC
CCAGGCTCCC 4551 CAGCAGGCAG AAGTATGCAA AGCATGCATC
TCAATTAGTC AGCAACCATA 4601 GTCCCGCCCC TAACTCCGCC CATCCCGCCC
CTAACTCCGC CCAGTTCCGC 4651 CCATTCTCCG CCCCATGGCT GACTAATTTT
TTTTATTTAT GCAGAGGCCG 4701 AGGCCGCCTC TGCCTCTGAG CTATTCCAGA
AGTAGTGAGG AGGCTTTTTT 4751 GGAGGCCTAG GCTTTTGCAA
AAAGCTCCCG GGAGCTTGTA TATCCATTTT 4801 CGGATCTGAT CAAGAGACAG
GATGAGGATC GTTTCGCATG ATTGAACAAG 4851 ATGGATTGCA CGCAGGTTCT
CCGGCCGCTT GGGTGGAGAG GCTATTCGGC 4901 TATGACTGGG
CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTTCCG 4951
GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT TGTCAAGACC GACCTGTCCG
5001 GTGCCCTGAA TGAACTGCAG GACGAGGCAG CGCGGCTATC
GTGGCTGGCC 5051 ACGACGGGCG TTCCTTGCGC AGCTGTGCTC
GACGTTGTCA CTGAAGCGGG 5101 AAGGGACTGG CTGCTATTGG
GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT 5151 CTCACCTTGC
TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG 5201
CGGCTGCATA CGCTTGATCC GGCTACCTGC CCATTCGACC ACCAAGCGAA 5251
ACATCGCATC GAGCGAGCAC GTACTCGGAT GGAAGCCGGT CTTGTCGATC
5301 AGGATGATCT GGACGAAGAG CATCAGGGGC TCGCGCCAGC
CGAACTGTTC 5351 GCCAGGCTCA AGGCGCGCAT GCCCGACGGC

GAGGATCTCG TCGTGACCCA 5401 TGGCGATGCC TGCTTGCCGA ATATCATGGT
GGAAAATGGC CGCTTTTCTG 5451 GATTCATCGA CTGTGGCCGG
CTGGGTGTGG CGGACCGCTA TCAGGACATA 5501 GCGTTGGCTA CCCGTGATAT
TGCTGAAGAG CTTGGCGGCG AATGGGCTGA 5551 CCGCTTCCTC
GTGCTTTACG GTATCGCCGC TCCCGATTCG CAGCGCATCG 5601 CCTTCTATCG
CCTTCTTGAC GAGTTCTTCT GAGCGGGACT CTGGGGTTCG 5651
AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGAGA TTTCGATTCC
5701 ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC
5751 CGGCTGGATG ATCCTCCAGC GCGGGGATCT CATGCTGGAG TTCTTCGCCC
5801 ACCCCAACTT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC
5851 ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT CTAGTTGTGG
5901 TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGTATA CCGTCGACCT
5951 CTAGCTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
6001 TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG
6051 TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC
6101 GCTCACTGCC CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA
6151 TGAATCGGCC AACGCGCGGG GAGAGGCGGT TTGCGTATTG
GGCGCTCTTC 6201 CGCTTCCTCG CTCACTGACT CGCTGCGCTC
GGTCGTTCGG CTGCGGCGAG 6251 CGGTATCAGC TCACTCAAAG
GCGGTAATAC GGTTATCCAC AGAATCAGGG 6301 GATAACGCAG
GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA 6351
CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG 6401
ACGAGCATCA CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA
6451 GGACTATAAA GATACCAGGC GTTTCCCCCT GGAAGCTCCC TCGTGCGCTC
6501 TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC TTTCTCCCTT
6551 CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG
6601 GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC
CCCCCGTTCA 6651 GCCCGACCGC TGCGCCTTAT CCGGTAACTA TCGTCTTGAG
TCCAACCCGG 6701 TAAGACACGA CTTATCGCCA CTGGCAGCAG
CCACTGGTAA CAGGATTAGC 6751 AGAGCGAGGT ATGTAGGCGG
TGCTACAGAG TTCTTGAAGT GGTGGCCTAA 6801 CTACGGCTAC
ACTAGAAGAA CAGTATTTGG TATCTGCGCT CTGCTGAAGC 6851 CAGTTACCTT
CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC 6901
ACCGCTGGTA GCGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA 6951
AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC 7001
AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA 7051
AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT 7101
CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA 7151
GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC 7201
TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG 7251
CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT 7301
TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
7351 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG
7401 AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA
7451 CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC
7501 GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
TGTGCAAAAA 7551 AGCGGTTAGC TCCTTCGGTC CTCCGATCGT
TGTCAGAAGT AAGTTGGCCG 7601 CAGTGTTATC ACTCATGGTT
ATGGCAGCAC TGCATAATTC TCTTACTGTC 7651 ATGCCATCCG TAAGATGCTT
TTCTGTGACT GGTGAGTACT CAACCAAGTC 7701 ATTCTGAGAA TAGTGTATGC

GGCGACCGAG TTGCTCTTGC CCGGCGTCAA 7751 TACGGGATAA
TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT 7801
GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG
7851 ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
7901 TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
GCAAAATGCC 7951 GCAAAAAAGG GAATAAGGGC GACACGGAAA
TGTTGAATAC TCATACTCTT 8001 CCTTTTTCAA TATTATTGAA GCATTTATCA
GGGTTATTGT CTCATGAGCG 8051 GATACATATT TGAATGTATT TAGAAAAATA
AACAAATAGG GGTTCCGCGC 8101 ACATTTCCCC GAAAAGTGCC ACCTGACGTC

(87) According to one embodiment, the vector is a plasmid and has the sequence of pUbC-(mNeonGreen).sub.4-tDeg (SEQ ID NO: 77; GenBank Accession No. MN052907.1, which is hereby incorporated by reference) as follows:

(88) TABLE-US-00011 1 GACGGATCGG GAGATCTCCC GATCCCCTAT
GGTGCACTCT CAGTACAATC 51 TGCTCTGATG CCGCATAGTT
AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT 101 GGAGGTCGCT
GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG 151
GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG
201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGGCCTC
CGCGCCGGGT 251 TTTGGCGCCT CCCGCGGGCG CCCCCCTCCT
CACGGCGAGC GCTGCCACGT 301 CAGACGAAGG GCGCAGCGAG
CGTCCTGATC CTTCCGCCCG GACGCTCAGG 351 ACAGCGGCCC
GCTGCTCATA AGACTCGGCC TTAGAACCCC AGTATCAGCA 401
GAAGGACATT TTAGGACGGG ACTTGGGTGA CTCTAGGGCA CTGGTTTTCT
451 TTCCAGAGAG CGGAACAGGC GAGGAAAAGT AGTCCCTTCT
CGGCGATTCT 501 GCGGAGGGAT CTCCGTGGGG CGGTGAACGC
CGATGATTAT ATAAGGACGC 551 GCCGGGTGTG GCACAGCTAG
TTCCGTCGCA GCCGGGATTT GGGTCGCGGT 601 TCTTGTTTGT
GGATCGCTGT GATCGTCACT TGGAAGCTTG CCACCATGGT 651
GAGCAAGGGC GAGGAGGATA ACATGGCCTC TCTCCCAGCG ACACATGAGT
701 TACACATCTT TGGCTCCATC AACGGTGTGG ACTTTGACAT
GGTGGGTCAG 751 GGCACCGGCA ATCCAAATGA TGGTTATGAG
GAGTTAAACC TGAAGTCCAC 801 CAAGGGTGAC CTCCAGTTCT
CCCCCTGGAT TCTGGTCCCT CATATCGGGT 851 ATGGCTTCCA
TCAGTACCTG CCCTACCCTG ACGGGATGTC GCCTTTCCAG 901
GCCGCCATGG TAGATGGCTC CGGATACCAA GTCCATCGCA CAATGCAGTT
951 TGAAGATGGT GCCTCCCTTA CTGTTAACTA CCGCTACACC
TACGAGGGAA 1001 GCCACATCAA AGGAGAGGCC CAGGTGAAGG
GGACTGGTTT CCCTGCTGAC 1051 GGTCCTGTGA TGACCAACTC
GCTGACCGCT GCGGACTGGT GCAGGTCGAA 1101 GAAGACTTAC
CCCAACGACA AAACCATCAT CAGTACCTTT AAGTGGAGTT 1151
ACACCACTGG AAATGGCAAG CGCTACCGGA GCACTGCGCG GACCACCTAC
1201 ACCTTTGCCA AGCCAATGGC GGCTAACTAT CTGAAGAACC
AGCCGATGTA 1251 CGTGTTCCGT AAGACGGAGC TCAAGCACTC
CAAGACCGAG CTCAACTTCA 1301 AGGAGTGGCA AAAGGCCTTT
ACCGATGTGA TGGGCATGGA CGAGCTGTAC 1351 AAGGGTGGAC
ATATGGGCAC AGGGTCCACA GGCGGTACCG GCGGAGTTTC 1401
CAAAGGAGAA GAAGACAATA TGGCATCACT CCCCGCAACC CACGAGTTGC
1451 ATATTTTCGG TTCAATTAAT GGAGTAGATT TCGATATGGT
TGGCCAGGGA 1501 ACAGGAAACC CAAACGACGG ATATGAAGAG
CTTAATCTCA AAAGTACCAA 1551 AGGCGATCTG CAATTTTCTC

CGTGGATACT CGTGCCACAC ATTGGATACG 1601 GATTTCACCA
ATATCTCCCG TATCCGGATG GAATGTCCCC CTTTCAAGCA 1651
GCAATGGTGG ACGGGAGTGG TTATCAGGTA CACAGAACCA TGCAGTTCGA
1701 GGACGGGGCT TCTCTGACCG TAAATTATAG GTATACTTAT
GAAGGCTCAC 1751 ATATTAAGGG CGAAGCACAG GTTAAAGGAA
CCGGGTTTCC TGCGGATGGC 1801 CCCGTCATGA CTAATTCTCT
GACAGCCGCA GATTGGTGTC GCTCCAAAAA 1851 GACATACCCG
AATGATAAGA CTATAATCTC AACATTCAAA TGGTCCTATA 1901
CGACAGGCAA CGGGAAACGA TATAGATCCA CGGCTCGAAC AACTTACACA
1951 TTCGCTAAAC CTATGGCCGC CAATTACCTC AAAAATCAGC
CCATGTATGT 2001 GTTTAGGAAA ACCGAATTGA AGCATTCTAA
AACGGAACTT AATTTTAAGG 2051 AATGGCAGAA GGCTTTCACA
GACGTAATGG GGATGGATGA ACTCTATAAA 2101 TCAGGTCTCG
AGTCCTCAGG GGGAACGGGT GGGTCCGGAG GAGTTAGTAA 2151
AGGTGAAGAG GACAATATGG CAAGTTTGCC TGCGACTCAC GAGCTTCATA
2201 TCTTTGGGTC TATAAATGGC GTTGACTTCG ATATGGTTGG
CCAAGGTACT 2251 GGCAACCCCA ATGACGGTTA CGAGGAGTTG
AATCTCAAGT CCACAAAAGG 2301 TGATCTTCAG TTCAGCCCTT
GGATTCTCGT ACCTCATATT GGATATGGCT 2351 TTCACCAGTA
CCTTCCATAC CCAGACGGTA TGTCACCCTT TCAAGCTGCG 2401
ATGGTGGATG GTTCCGGCTA TCAGGTCCAC CGAACGATGC AATTCGAGGA
2451 CGGGGCCAGC CTCACCGTTA ATTATAGGTA CACCTATGAG
GGAAGTCACA 2501 TAAAGGGAGA AGCCCAAGTG AAAGGAACAG
GATTCCCAGC TGATGGTCCA 2551 GTAATGACGA ACTCCTTGAC
AGCGGCTGAC TGGTGTAGAA GCAAAAAGAC 2601 GTATCCTAAT
GACAAGACCA TCATTAGCAC TTTCAAATGG AGTTATACCA 2651
CAGGAAACGG CAAACGGTAC AGAAGCACTG CTAGAACTAC CTACACTTTC
2701 GCAAAGCCGA TGGCTGCAAA CTATTTGAAG AATCAGCCCA
TGTACGTGTT 2751 TCGAAAAACG GAACTTAAGC ACAGTAAGAC
TGAACTTAAT TTCAAGGAGT 2801 GGCAGAAGGC GTTCACGGAT
GTCATGGGTA TGGATGAACT GTATAAGGGA 2851 GGGTCTGGCA
CTGGGGGCAC TGCCAGCAGC GGATCCGGTG GCGGTGTGAG 2901
CAAGGGCGAG GAGGATAACA TGGCCTCTCT CCCAGCGACA CATGAGTTAC
2951 ACATCTTTGG CTCCATCAAC GGTGTGGACT TTGACATGGT
GGGTCAGGGC 3001 ACCGGCAATC CAAATGATGG TTATGAGGAG
TTAAACCTGA AGTCCACCAA 3051 GGGTGACCTC CAGTTCTCCC
CCTGGATTCT GGTCCCTCAT ATCGGGTATG 3101 GCTTCCATCA
GTACCTGCCC TACCCTGACG GGATGTCGCC TTTCCAGGCC 3151
GCCATGGTAG ATGGCTCCGG ATACCAAGTC CATCGCACAA TGCAGTTTGA
3201 AGATGGTGCC TCCCTTACTG TTAACTACCG CTACACCTAC
GAGGGAAGCC 3251 ACATCAAAGG AGAGGCCCAG GTGAAGGGGA
CTGGTTTCCC TGCTGACGGT 3301 CCTGTGATGA CCAACTCGCT
GACCGCTGCG GACTGGTGCA GGTCGAAGAA 3351 GACTTACCCC
AACGACAAAA CCATCATCAG TACCTTTAAG TGGAGTTACA 3401
CCACTGGAAA TGGCAAGCGC TACCGGAGCA CTGCGCGGAC CACCTACACC
3451 TTTGCCAAGC CAATGGCGGC TAACTATCTG AAGAACCAGC
CGATGTACGT 3501 GTTCCGTAAG ACGGAGCTCA AGCACTCCAA
GACCGAGCTC AACTTCAAGG 3551 AGTGGCAAAA GGCCTTTACC
GATGTGATGG GCATGGACGA GCTGTACAAG 3601 GGCGGAAGAT
CCGGTGGTGG TTCTGGTCCT CGTCCCCGTG GTACTCGTGG 3651

TAAAGGTCGC CGTATTCGTC GCCGCGGTTA ATCTAGAGGG CCCGTTTAAA
3701 CCCGCTGATC AGCCTCGACT GTGCCTTCTA GTTGCCAGCC
ATCTGTTGTT 3751 TGCCCCTCCC CCGTGCCTTC CTTGACCCTG
GAAAGGTGCC ACTCCCACTG 3801 TCCTTTCCTA ATAAAATGAG
GAAATTGCAT CGCATTGTCT GAGTAGGTGT 3851 CATTCTATTC
TGGGGGGTGG GGGTGGGGGC AGGACAGCAA GGGGGAGGAT 3901
TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGCTTC
3951 TGAGGCGGAA AGAACCAGCT GGGGCTCTAG GGGGTATCCC
CACGCGCCCT 4001 GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG
TGGTTACGCG CAGCGTGACC 4051 GCTACACTTG CCAGCGCCCT
AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC 4101 CTTTCTCGCC
ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA AATCGGGGGC 4151
TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAAAAAA
4201 CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT
GATAGACGGT 4251 TTTTCGCCCT TTGACGTTGG AGTCCACGTT
CTTTAATAGT GGACTCTTGT 4301 TCCAAACTGG AACAACACTC
AACCCTATCT CGGTCTATTC TTTTGATTTA 4351 TAAGGGATTT
TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA 4401
ACAAAAATTT AACGCGAATT AATTCTGTGG AATGTGTGTC AGTTAGGGTG
4451 TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA
AGCATGCATC 4501 TCAATTAGTC AGCAACCAGG TGTGGAAAGT
CCCCAGGCTC CCCAGCAGGC 4551 AGAAGTATGC AAAGCATGCA
TCTCAATTAG TCAGCAACCA TAGTCCCGCC 4601 CCTAACTCCG
CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC 4651
CGCCCCATGG CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC
4701 TCTGCCTCTG AGCTATTCCA GAAGTAGTGA GGAGGCTTTT
TTGGAGGCCT 4751 AGGCTTTTGC AAAAAGCTCC CGGGAGCTTG
TATATCCATT TTCGGATCTG 4801 ATCAAGAGAC AGGATGAGGA
TCGTTTCGCA TGATTGAACA AGATGGATTG 4851 CACGCAGGTT
CTCCGGCCGC TTGGGTGGAG AGGCTATTCG GCTATGACTG 4901
GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG
4951 CGCAGGGGCG CCCGGTTCTT TTTGTCAAGA CCGACCTGTC
CGGTGCCCTG 5001 AATGAACTGC AGGACGAGGC AGCGCGGCTA
TCGTGGCTGG CCACGACGGG 5051 CGTTCCTTGC GCAGCTGTGC
TCGACGTTGT CACTGAAGCG GGAAGGGACT 5101 GGCTGCTATT
GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC ATCTCACCTT 5151
GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA
5201 TACGCTTGAT CCGGCTACCT GCCCATTCGA CCACCAAGCG
AAACATCGCA 5251 TCGAGCGAGC ACGTACTCGG ATGGAAGCCG
GTCTTGTCGA TCAGGATGAT 5301 CTGGACGAAG AGCATCAGGG
GCTCGCGCCA GCCGAACTGT TCGCCAGGCT 5351 CAAGGCGCGC
ATGCCCGACG GCGAGGATCT CGTCGTGACC CATGGCGATG 5401
CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTCATC
5451 GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA
TAGCGTTGGC 5501 TACCCGTGAT ATTGCTGAAG AGCTTGGCGG
CGAATGGGCT GACCGCTTCC 5551 TCGTGCTTTA CGGTATCGCC
GCTCCCGATT CGCAGCGCAT CGCCTTCTAT 5601 CGCCTTCTTG
ACGAGTTCTT CTGAGCGGGA CTCTGGGGTT CGAAATGACC 5651
GACCAAGCGA CGCCCAACCT GCCATCACGA GATTTCGATT CCACCGCCGC
5701 CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC

GCCGGCTGGA 5751 TGATCCTCCA GCGCGGGGAT CTCATGCTGG
AGTTCTTCGC CCACCCCAAC 5801 TTGTTTATTG CAGCTTATAA
TGGTTACAAA TAAAGCAATA GCATCACAAA 5851 TTTCACAAAT
AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA 5901
AACTCATCAA TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG
5951 AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA
ATTGTTATCC 6001 GCTCACAATT CCACACAACA TACGAGCCGG
AAGCATAAAG TGTAAAGCCT 6051 GGGGTGCCTA ATGAGTGAGC
TAACTCACAT TAATTGCGTT GCGCTCACTG 6101 CCCGCTTTCC
AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG 6151
CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT
6201 CGCTCACTGA CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG
AGCGGTATCA 6251 GCTCACTCAA AGGCGGTAAT ACGGTTATCC
ACAGAATCAG GGGATAACGC 6301 AGGAAAGAAC ATGTGAGCAA
AAGGCCAGCA AAAGGCCAGG AACCGTAAAA 6351 AGGCCGCGTT
GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT 6401
CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA
6451 AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC
TCTCCTGTTC 6501 CGACCCTGCC GCTTACCGGA TACCTGTCCG
CCTTTCTCCC TTCGGGAAGC 6551 GTGGCGCTTT CTCATAGCTC
ACGCTGTAGG TATCTCAGTT CGGTGTAGGT 6601 CGTTCGCTCC
AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC 6651
GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC
6701 GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA
GCAGAGCGAG 6751 GTATGTAGGC GGTGCTACAG AGTTCTTGAA
GTGGTGGCCT AACTACGGCT 6801 ACACTAGAAG AACAGTATTT
GGTATCTGCG CTCTGCTGAA GCCAGTTACC 6851 TTCGGAAAAA
GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG 6901
TAGCGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT
6951 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC
TCAGTGGAAC 7001 GAAAACTCAC GTTAAGGGAT TTTGGTCATG
AGATTATCAA AAAGGATCTT 7051 CACCTAGATC CTTTTAAATT
AAAAATGAAG TTTTAAATCA ATCTAAAGTA 7101 TATATGAGTA
AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA 7151
CCTATCTCAG CGATCTGTCT ATTTCGTTCA TCCATAGTTG CCTGACTCCC
7201 CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT
GGCCCCAGTG 7251 CTGCAATGAT ACCGCGAGAC CCACGCTCAC
CGGCTCCAGA TTTATCAGCA 7301 ATAAACCAGC CAGCCGGAAG
GGCCGAGCGC AGAAGTGGTC CTGCAACTTT 7351 ATCCGCCTCC
ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA 7401
GTTCGCCAGT TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC
7451 GTGGTGTCAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT
CCGGTTCCCA 7501 ACGATCAAGG CGAGTTACAT GATCCCCCAT
GTTGTGCAAA AAAGCGGTTA 7551 GCTCCTTCGG TCCTCCGATC
GTTGTCAGAA GTAAGTTGGC CGCAGTGTTA 7601 TCACTCATGG
TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC 7651
CGTAAGATGC TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG
7701 AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC
AATACGGGAT 7751 AATACCGCGC CACATAGCAG AACTTTAAAA
GTGCTCATCA TTGGAAAACG 7801 TTCTTCGGGG CGAAAACTCT

CAAGGATCTT ACCGCTGTTG AGATCCAGTT 7851 CGATGTAACC
CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC 7901
ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA
7951 GGGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC
TTCCTTTTTC 8001 AATATTATTG AAGCATTTAT CAGGGTTATT
GTCTCATGAG CGGATACATA 8051 TTTGAATGTA TTTAGAAAAA
TAAACAAATA GGGGTTCCGC GCACATTTCC 8101 CCGAAAAGTG
CCACCTGACG TC

(89) According to one embodiment, the vector is a plasmid and has the sequence of pAV-U6+27-Tornado-F30-Pepper(TAR Variant-2) (SEQ ID NO: 78; GenBank Accession No. MN052908.1, which is hereby incorporated by reference in its entirety) as follows:

(90) TABLE-US-00012 1 GCCGGATCCA AGGTCGGGCA GGAAGAGGGC
CTATTTCCCA TGATTCCTTC 51 ATATTTGCAT ATACGATACA
AGGCTGTTAG AGAGATAATT AGAATTAATT 101 TGACTGTAAA
CACAAAGATA TTAGTACAAA ATACGTGACG TAGAAAGTAA 151
TAATTTCTTG GGTAGTTTGC AGTTTTAAAA TTATGTTTTA AAATGGACTA
201 TCATATGCTT ACCGTAACTT GAAAGTATTT CGATTTCTTG GCTTTATATA
251 TCTTGTGGAA AGGACGAAAC ACCGTGCTCG CTTCGGCAGC
ACATATACTA 301 GTCGACGGGC CGCACTCGCC GGTCCCAAGC
CCGGATAAAA TGGGAGGGGG 351 CGGGAAACCG CCTAACCATG
CCGAGTGCGG CCGCTTGCCA TGTGTATGTG 401 GGACGCGTTG
CCACGTTTCC CACATACTCT GATGATCCGC TAGCAAAGGC 451
TCGTTGAGCT CATTAGCTCC GAGCCCGAGG TACCGGATCA TTCATGGCAA
501 GCGGCCGCGG TCGGCGTGGA CTGTAGAACA CTGCCAATGC
CGGTCCCAAG 551 CCCGGATAAA AGTGGAGGGT ACAGTCCACG
CTCTAGAGCG GACTTCGGTC 601 CGCTTTTTAC TAGGACCTGC
AGGCATGCAA GCTTGACGTC GGTTACCGAT 651 ATCCATATGG
CGACCGCATC GATCTCGAGC CGAGGACTAG TAACTTGTTT 701
ATTGCAGCTT ATAATGGTTA CAAATAAAGC AATAGCATCA CAAATTTCAC
751 AAATAAAGCA TTTTTTTCAC TGCATTCTAG TTGTGGTTTG
TCCAAACTCA 801 TCAATGTATC TTATCATGTC TTACGTAGAT
AAGTAGCATG GCGGGTTAAT 851 CATTAACTAC AAGGAACCCC
TAGTGATGGA GTTGGCCACT CCCTCTCTGC 901 GCGCTCGCTC
GCTCACTGAG GCCGGGCGAC CAAAGGTCGC CCGACGCCCG 951
GGCTTTGCCC GGGCGGCCTC AGTGAGCGAG CGAGCGCGCA GAGAGGGAGT
1001 GGCCAAAGAT CTCTGGCGTA ATAGCGAAGA GGCCCGCACC
GATCGCCCTT 1051 CCCAACAGTT GCGCAGCCTG AATGGCTAAT
GGGAAATTGT AAACGTTAAT 1101 ATTTTGTTAA TATTTTGTTA
AAATTCGCGT TAAATTTTTG TTAAATCAGC 1151 TCATTTTTTA
ACCAATAGGC CGAAATCGGC AAAATCCCTT ATAAATCAAA 1201
AGAATAGACC GAGATAGGGT TGAGTGTTGT TCCAGTTTGG AACAAGAGTC
1251 CACTATTAAA GAACGTGGAC TCCAACGTCA AAGGGCGAAA
AACCGTCTAT 1301 CAGGGCGATG GCCCACTACG TGAACCATCA
CCCTAATCAA GTTTTTTGGG 1351 GTCGAGGTGC CGTAAAGCAC
TAAATCGGAA CCCTAAAGGG ATGCCCCGAT 1401 TTAGAGCTTG
ACGGGGAAAG CCGGCGAACG TGGCGAGAAA GGAAGGGAAG 1451
AAAGCGAAAG GAGCGGGCGC TAGGGCGCTG GCAAGTGTAG CGGTCACGCT
1501 GCGCGTAACC ACCACACCCG CCGCGCTTAA TGCGCCGCTA
CAGGGCGCGT 1551 CAGGTGGCAC TTTTCGGGGA AATGTGCGCG
GAACCCCTAT TTGTTTATTT 1601 TTCTAAATAC ATTCAAATAT

GTATCCGCTC ATGAGACAAT AACCCTGATA 1651 AATGCTTCAA
TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC 1701
GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT
1751 CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
AGTTGGGTGC 1801 ACGAGTGGGT TACATCGAAC TGGATCTCAA
CAGCGGTAAG ATCCTTGAGA 1851 GTTTTCGCCC CGAAGAACGT
TTTCCAATGA TGAGCACTTT TAAAGTTCTG 1901 CTATGTGGCG
CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG 1951
TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA
2001 CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
ATGCAGTGCT 2051 GCCATAACCA TGAGTGATAA CACTGCGGCC
AACTTACTTC TGACAACGAT 2101 CGGAGGACCG AAGGAGCTAA
CCGCTTTTTT GCACAACATG GGGGATCATG 2151 TAACTCGCCT
TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC 2201
GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2251 ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA
CAATTAATAG 2301 ACTGGATGGA GGCGGATAAA GTTGCAGGAC
CACTTCTGCG CTCGGCCCTT 2351 CCGGCTGGCT GGTTTATTGC
TGATAAATCT GGAGCCGGTG AGCGTGGGTC 2401 TCGCGGTATC
ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG 2451
TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA
2501 CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT
AACTGTCAGA 2551 CCAAGTTTAC TCATATATAC TTTAGATTGA
TTTAAAACTT CATTTTTAAT 2601 TTAAAAGGAT CTAGGTGAAG
ATCCTTTTTG ATAATCTCAT GACCAAAATC 2651 CCTTAACGTG
AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT 2701
CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC
2751 AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
GGATCAAGAG 2801 CTACCAACTC TTTTTCCGAA GGTAACTGGC
TTCAGCAGAG CGCAGATACC 2851 AAATACTGTC CTTCTAGTGT
AGCCGTAGTT AGGCCACCAC TTCAAGAACT 2901 CTGTAGCACC
GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT 2951
GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA
3001 GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
TCGTGCAACA 3051 CAGCCAGCTT GGAGCGAACG ACCTACACCG
AACTGAGATA CCTACAGCGT 3101 GAGCATTGAG AAAGCGCCAC
GCTTCCCGAA GGGAGAAAGG CGGACAGGTA 3151 TCCGGTAAGC
GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG 3201
GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
3251 CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
GCCTATGGAA 3301 AAACGCCAGC AACGCGGCCT TTTTACGGTT
CCTGGCCTTT TGCTGGCCTT 3351 TTGCTCACAT GTTCTTTCCT
GCGTTATCCC CTGATTCTGT GGATAACCGT 3401 ATTACCGCCT
TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA 3451
GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC
3501 CGCCTCTCCC CGCGCGTTGG CCGATTCATT AATGCAGAGA
TCTTTGGCCA 3551 CTCCCTCTCT GCGCGCTCGC TCGCTCACTG
AGGCCGGGCG ACCAAAGGTC 3601 GCCCGACGCC CGGGCTTTGC
CCGGGCGGCC TCAGTGAGCG AGCGAGCGCG 3651 CAGAGAGGGA
GTGGCCAACT CCATCACTAG GGGTTCCTGG AGGGGTGGAG 3701

TCGTGACGTG AATTACGTCA TAGGGTTAGG GAGGTCCTGG ATCGATCCAG
3751 ACATGATAAG ATACATTGAT GAGTTTGGAC AAACCACAAC
TAGAATGCAG 3801 TGAAAAAAAT GCTTTATTTG TGAAATTTGT
GATGCTATTG CTTTATTTGT 3851 AACCATTATA AGCTGCAATA
AACAAGTTAA CAACAACAAT TGCATTCATT 3901 TTATGTTTCA
GGTTCAGGGG GAGGTGTGGG AGGTTTTTTA AAGCAAGTAA 3951
AACCTCTACA AATGTGGTAT GGCTGATTAT GATCTCTAGT CAAGGCACTA
4001 TACATCAAAT ATTCCTTATT AACCCCTTTA CAAATTAAAA
AGCTAAAGGT 4051 ACACAATTTT TGAGCATAGT TATTAATAGC
AGACACTCTA TGCCTGTGTG 4101 GAGTAAGAAA AAACAGTATG
TTATGATTAT AACTGTTATG CCTACTTATA 4151 AAGGTTACAG
AATATTTTTC CATAATTTTC TTGTATAGCA GTGCAGCTTT 4201
TTCCTTTGTG GTGTAAATAG CAAAGCAAGC AAGAGTTCTA TTACTAAACA
4251 CAGCATGACT CAAAAAACTT AGCAATTCTG AAGGAAAGTC
CTTGGGGTCT 4301 TCTACCTTTC TCTTCTTTTT TGGAGGAGTA
GAATGTTGAG AGTCAGCAGT 4351 AGCCTCATCA TCACTAGATG
GCATTTCTTC TGAGCAAAAC AGGTTTTCCT 4401 CATTAAAGGC
ATTCCACCAC TGCTCCCATT CATCAGTTCC ATAGGTTGGA 4451
ATCTAAAATA CACAAACAAT TAGAATCAGT AGTTTAACAC ATTATACACT
4501 TAAAAATTTT ATATTTACCT TAGAGCTTTA AATCTCTGTA
GGTAGTTTGT 4551 CCAATTATGT CACACCACAG AAGTAAGGTT
CCTTCACAAA GATCCGGGAC 4601 CAAAGCGGCC ATCGTGCCTC
CCCACTCCTG CAGTTCGGGG GCATGGATGC 4651 GCGGATAGCC
GCTGCTGGTT TCCTGGATGC CGACGGATTT GCACTGCCGG 4701
TAGAACTCCG CGAGGTCGTC CAGCCTCAGG CAGCAGCTGA ACCAACTCGC
4751 GAGGGGATCG AGCCCGGGGT GGGCGAAGAA CTCCAGCATG
AGATCCCCGC 4801 GCTGGAGGAT CATCCAGCCG GCGTCCCGGA
AAACGATTCC GAAGCCCAAC 4851 CTTTCATAGA AGGCGGCGGT
GGAATCGAAA TCTCGTGATG GCAGGTTGGG 4901 CGTCGCTTGG
TCGGTCATTT CGAACCCCAG AGTCCCGCTC AGAAGAACTC 4951
GTCAAGAAGG CGATAGAAGG CGATGCGCTG CGAATCGGGA GCGGCGATAC
5001 CGTAAAGCAC GAGGAAGCGG TCAGCCCATT CGCCGCCAAG
CTCTTCAGCA 5051 ATATCACGGG TAGCCAACGC TATGTCCTGA
TAGCGGTCCG CCACACCCAG 5101 CCGGCCACAG TCGATGAATC
CAGAAAAGCG GCCATTTTCC ACCATGATAT 5151 TCGGCAAGCA
GGCATCGCCA TGGGTCACGA CGAGATCCTC GCCGTCGGGC 5201
ATGCGCGCCT TGAGCCTGGC GAACAGTTCG GCTGGCGCGA GCCCCTGATG
5251 CTCTTGTCCA GATCATCCTG ATCGACAAGA CCGGCTTCCA
TCCGAGTACG 5301 TGCTCGCTCG ATGCGATGTT CGCTTGGTGG
TCGAATGGGC AGGTAGCCGG 5351 ATCAAGCGTA TGCAGCCGCC
GCATTGCATC AGCCATGATG GATACTTTCT 5401 CGGCAGGAGC
AAGGTGAGAT GACAGGAGAT CCTGCCCCGG CACTTCGCCC 5451
AATAGCAGCC AGTCCCTTCC CGCTTCAGTG ACAACGTCGA GCACAGCTGC
5501 GCAAGGAACG CCCGTCGTGG CCAGCCACGA TAGCCGCGCT
GCCTCGTCCT 5551 GCAGTTCATT CAGGGCACCG GACAGGTCGG
TCTTGACAAA AAGAACCGGG 5601 CGCCCCTGCG CTGACAGCCG
GAACACGGCG GCATCAGAGC AGCCGATTGT 5651 CTGTTGTGCC
CAGTCATAGC CGAATAGCCT CTCCACCCAA GCGGCCGGAG 5701
AACCTGCGTG CAATCCATCT TGTTCAATCA TGCGAAACGA TCCTCATCCT
5751 GTCTCTTGAT CAGATCTTGA TCCCCTGCGC CATCAGATCC

TTGGCGGCAA 5801 GAAAGCCATC CAGTTTACTT TGCAGGGCTT
CCCAACCTTA CCAGAGGGCG 5851 CCCCAGCTGG CAATTCCGGT
TCGCTTGCTG TCCATAAAAC CGCCCAGTCT 5901 AGCTATCGGC
ATGTAAGCCC ACTGCAAGCT ACCTGCTTTC TCTTTGCGCT 5951
TGCGTTTTCC CTTGTCCAGA TAGCCCAGTA GCTGACATTC ATCCGGGGTC
6001 AGCACCGTTT CTGCGGACTG GCTTTCTACG TGTTCCGCTT
CCTTTAGCAG 6051 CCCTTGCGCC CTGAGTGCTT GCGGCAGCGT
GAAGCTTTTT GCAAAAGCCT 6101 AGGCCTCCAA AAAAGCCTCC
TCACTACTTC TGGAATAGCT CAGAGGCCGA 6151 GGCGGCCTCG
GCCTCTGCAT AAATAAAAAA AATTAGTCAG CCATGGGGCG 6201
GAGAATGGGC GGAACTGGGC GGAGTTAGGG GCGGGATGGG CGGAGTTAGG
6251 GGCGGGACTA TGGTTGCTGA CTAATTGAGA TGCATGCTTT
GCATACTTCT 6301 GCCTGCTGGG GAGCCTGGGG ACTTTCCACA
CCTGGTTGCT GACTAATTGA 6351 GATGCATGCT TTGCATACTT
CTGCCTGCTG GGGAGCCTGG GGACTTTCCA 6401 CACCCTAACT
GACACACATT CCACA

(91) According to one embodiment, the vector is a plasmid and has the sequence of pAV-U6+27-Tornado-F30-TAR Variant-1 (SEQ ID NO: 79; GenBank Accession No. MN052909.1, which is hereby incorporated by reference in its entirety) as follows:

(92) TABLE-US-00013 1 GCCGGATCCA AGGTCGGGCA GGAAGAGGGC
CTATTTCCCA TGATTCCTTC 51 ATATTTGCAT ATACGATACA
AGGCTGTTAG AGAGATAATT AGAATTAATT 101 TGACTGTAAA
CACAAAGATA TTAGTACAAA ATACGTGACG TAGAAAGTAA 151
TAATTTCTTG GGTAGTTTGC AGTTTTAAAA TTATGTTTTA AAATGGACTA
201 TCATATGCTT ACCGTAACTT GAAAGTATTT CGATTTCTTG GCTTTATATA
251 TCTTGTGGAA AGGACGAAAC ACCGTGCTCG CTTCGGCAGC
ACATATACTA 301 GTCGACGGGC CGCACTCGCC GGTCCCAAGC
CCGGATAAAA TGGGAGGGGG 351 CGGGAAACCG CCTAACCATG
CCGAGTGCGG CCGCTTGCCA TGTGTATGTG 401 GGACGCGTTG
CCACGTTTCC CACATACTCT GATGATCCGC TAGCAAAGGC 451
TCGTCTGAGC TCATTAGCTC CGAGCCCGAG GTACCGGATC ATTCATGGCA
501 AGCGGCCGCG GTCGGCGTGG ACTGTAGAAC ACTGCCAATG
CCGGTCCCAA 551 GCCCGGATAA AAGTGGAGGG TACAGTCCAC
GCTCTAGAGC GGACTTCGGT 601 CCGCTTTTTA CTAGGACCTG
CAGGCATGCA AGCTTGACGT CGGTTACCGA 651 TATCCATATG
GCGACCGCAT CGATCTCGAG CCGAGGACTA GTAACTTGTT 701
TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA
751 CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT
GTCCAAACTC 801 ATCAATGTAT CTTATCATGT CTTACGTAGA
TAAGTAGCAT GGCGGGTTAA 851 TCATTAACTA CAAGGAACCC
CTAGTGATGG AGTTGGCCAC TCCCTCTCTG 901 CGCGCTCGCT
CGCTCACTGA GGCCGGGCGA CCAAAGGTCG CCCGACGCCC 951
GGGCTTTGCC CGGGCGGCCT CAGTGAGCGA GCGAGCGCGC AGAGAGGGAG
1001 TGGCCAAAGA TCTCTGGCGT AATAGCGAAG AGGCCCGCAC
CGATCGCCCT 1051 TCCCAACAGT TGCGCAGCCT GAATGGCTAA
TGGGAAATTG TAAACGTTAA 1101 TATTTTGTTA ATATTTTGTT
AAAATTCGCG TTAAATTTTT GTTAAATCAG 1151 CTCATTTTTT
AACCAATAGG CCGAAATCGG CAAAATCCCT TATAAATCAA 1201
AAGAATAGAC CGAGATAGGG TTGAGTGTTG TTCCAGTTTG GAACAAGAGT
1251 CCACTATTAA AGAACGTGGA CTCCAACGTC AAAGGGCGAA

AAACCGTCTA 1301 TCAGGGCGAT GGCCCACTAC GTGAACCATC
ACCCTAATCA AGTTTTTTGG 1351 GGTCGAGGTG CCGTAAAGCA
CTAAATCGGA ACCCTAAAGG GATGCCCCGA 1401 TTTAGAGCTT
GACGGGGAAA GCCGGCGAAC GTGGCGAGAA AGGAAGGGAA 1451
GAAAGCGAAA GGAGCGGGCG CTAGGGCGCT GGCAAGTGTA GCGGTCACGC
1501 TGCGCGTAAC CACCACACCC GCCGCGCTTA ATGCGCCGCT
ACAGGGCGCG 1551 TCAGGTGGCA CTTTTCGGGG AAATGTGCGC
GGAACCCCTA TTTGTTTATT 1601 TTTCTAAATA CATTCAAATA
TGTATCCGCT CATGAGACAA TAACCCTGAT 1651 AAATGCTTCA
ATAATATTGA AAAAGGAAGA GTATGAGTAT TCAACATTTC 1701
CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC
1751 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT
CAGTTGGGTG 1801 CACGAGTGGG TTACATCGAA CTGGATCTCA
ACAGCGGTAA GATCCTTGAG 1851 AGTTTTCGCC CCGAAGAACG
TTTTCCAATG ATGAGCACTT TTAAAGTTCT 1901 GCTATGTGGC
GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG 1951
GTCGCCGCAT ACACTATTCT CAGAATGACT TGGTTGAGTA CTCACCAGTC
2001 ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT
TATGCAGTGC 2051 TGCCATAACC ATGAGTGATA ACACTGCGGC
CAACTTACTT CTGACAACGA 2101 TCGGAGGACC GAAGGAGCTA
ACCGCTTTTT TGCACAACAT GGGGGATCAT 2151 GTAACTCGCC
TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA 2201
CGACGAGCGT GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA
2251 AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA
ACAATTAATA 2301 GACTGGATGG AGGCGGATAA AGTTGCAGGA
CCACTTCTGC GCTCGGCCCT 2351 TCCGGCTGGC TGGTTTATTG
CTGATAAATC TGGAGCCGGT GAGCGTGGGT 2401 CTCGCGGTAT
CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC 2451
GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG
2501 ACAGATCGCT GAGATAGGTG CCTCACTGAT TAAGCATTGG
TAACTGTCAG 2551 ACCAAGTTTA CTCATATATA CTTTAGATTG
ATTTAAAACT TCATTTTTAA 2601 TTTAAAAGGA TCTAGGTGAA
GATCCTTTTT GATAATCTCA TGACCAAAAT 2651 CCCTTAACGT
GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA 2701
TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG
2751 CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC
CGGATCAAGA 2801 GCTACCAACT CTTTTTCCGA AGGTAACTGG
CTTCAGCAGA GCGCAGATAC 2851 CAAATACTGT CCTTCTAGTG
TAGCCGTAGT TAGGCCACCA CTTCAAGAAC 2901 TCTGTAGCAC
CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC 2951
TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT
3001 AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG
TTCGTGCAAC 3051 ACAGCCAGCT TGGAGCGAAC GACCTACACC
GAACTGAGAT ACCTACAGCG 3101 TGAGCATTGA GAAAGCGCCA
CGCTTCCCGA AGGGAGAAAG GCGGACAGGT 3151 ATCCGGTAAG
CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA 3201
GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
3251 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG
AGCCTATGGA 3301 AAAACGCCAG CAACGCGGCC TTTTTACGGT
TCCTGGCCTT TTGCTGGCCT 3351 TTTGCTCACA TGTTCTTTCC

TGCGTTATCC CCTGATTCTG TGGATAACCG 3401 TATTACCGCC
TTTGAGTGAG CTGATACCGC TCGCCGCAGC CGAACGACCG 3451
AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA
3501 CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGAG
ATCTTTGGCC 3551 ACTCCCTCTC TGCGCGCTCG CTCGCTCACT
GAGGCCGGGC GACCAAAGGT 3601 CGCCCGACGC CCGGGCTTTG
CCCGGGCGGC CTCAGTGAGC GAGCGAGCGC 3651 GCAGAGAGGG
AGTGGCCAAC TCCATCACTA GGGGTTCCTG GAGGGGTGGA 3701
GTCGTGACGT GAATTACGTC ATAGGGTTAG GGAGGTCCTG GATCGATCCA
3751 GACATGATAA GATACATTGA TGAGTTTGGA CAAACCACAA
CTAGAATGCA 3801 GTGAAAAAAA TGCTTTATTT GTGAAATTTG
TGATGCTATT GCTTTATTTG 3851 TAACCATTAT AAGCTGCAAT
AAACAAGTTA ACAACAACAA TTGCATTCAT 3901 TTTATGTTTC
AGGTTCAGGG GGAGGTGTGG GAGGTTTTTT AAAGCAAGTA 3951
AAACCTCTAC AAATGTGGTA TGGCTGATTA TGATCTCTAG TCAAGGCACT
4001 ATACATCAAA TATTCCTTAT TAACCCCTTT ACAAATTAAA
AAGCTAAAGG 4051 TACACAATTT TTGAGCATAG TTATTAATAG
CAGACACTCT ATGCCTGTGT 4101 GGAGTAAGAA AAAACAGTAT
GTTATGATTA TAACTGTTAT GCCTACTTAT 4151 AAAGGTTACA
GAATATTTTT CCATAATTTT CTTGTATAGC AGTGCAGCTT 4201
TTTCCTTTGT GGTGTAAATA GCAAAGCAAG CAAGAGTTCT ATTACTAAAC
4251 ACAGCATGAC TCAAAAAACT TAGCAATTCT GAAGGAAAGT
CCTTGGGGTC 4301 TTCTACCTTT CTCTTCTTTT TTGGAGGAGT
AGAATGTTGA GAGTCAGCAG 4351 TAGCCTCATC ATCACTAGAT
GGCATTTCTT CTGAGCAAAA CAGGTTTTCC 4401 TCATTAAAGG
CATTCCACCA CTGCTCCCAT TCATCAGTTC CATAGGTTGG 4451
AATCTAAAAT ACACAAACAA TTAGAATCAG TAGTTTAACA CATTATACAC
4501 TTAAAAATTT TATATTTACC TTAGAGCTTT AAATCTCTGT
AGGTAGTTTG 4551 TCCAATTATG TCACACCACA GAAGTAAGGT
TCCTTCACAA AGATCCGGGA 4601 CCAAAGCGGC CATCGTGCCT
CCCCACTCCT GCAGTTCGGG GGCATGGATG 4651 CGCGGATAGC
CGCTGCTGGT TTCCTGGATG CCGACGGATT TGCACTGCCG 4701
GTAGAACTCC GCGAGGTCGT CCAGCCTCAG GCAGCAGCTG AACCAACTCG
4751 CGAGGGGATC GAGCCCGGGG TGGGCGAAGA ACTCCAGCAT
GAGATCCCCG 4801 CGCTGGAGGA TCATCCAGCC GGCGTCCCGG
AAAACGATTC CGAAGCCCAA 4851 CCTTTCATAG AAGGCGGCGG
TGGAATCGAA ATCTCGTGAT GGCAGGTTGG 4901 GCGTCGCTTG
GTCGGTCATT TCGAACCCCA GAGTCCCGCT CAGAAGAACT 4951
CGTCAAGAAG GCGATAGAAG GCGATGCGCT GCGAATCGGG AGCGGCGATA
5001 CCGTAAAGCA CGAGGAAGCG GTCAGCCCAT TCGCCGCCAA
GCTCTTCAGC 5051 AATATCACGG GTAGCCAACG CTATGTCCTG
ATAGCGGTCC GCCACACCCA 5101 GCCGGCCACA GTCGATGAAT
CCAGAAAAGC GGCCATTTTC CACCATGATA 5151 TTCGGCAAGC
AGGCATCGCC ATGGGTCACG ACGAGATCCT CGCCGTCGGG 5201
CATGCGCGCC TTGAGCCTGG CGAACAGTTC GGCTGGCGCG AGCCCCTGAT
5251 GCTCTTGTCC AGATCATCCT GATCGACAAG ACCGGCTTCC
ATCCGAGTAC 5301 GTGCTCGCTC GATGCGATGT TCGCTTGGTG
GTCGAATGGG CAGGTAGCCG 5351 GATCAAGCGT ATGCAGCCGC
CGCATTGCAT CAGCCATGAT GGATACTTTC 5401 TCGGCAGGAG
CAAGGTGAGA TGACAGGAGA TCCTGCCCCG GCACTTCGCC 5451

CAATAGCAGC CAGTCCCTTC CCGCTTCAGT GACAACGTCG AGCACAGCTG
5501 CGCAAGGAAC GCCCGTCGTG GCCAGCCACG ATAGCCGCGC
TGCCTCGTCC 5551 TGCAGTTCAT TCAGGGCACC GGACAGGTCG
GTCTTGACAA AAAGAACCGG 5601 GCGCCCCTGC GCTGACAGCC
GGAACACGGC GGCATCAGAG CAGCCGATTG 5651 TCTGTTGTGC
CCAGTCATAG CCGAATAGCC TCTCCACCCA AGCGGCCGGA 5701
GAACCTGCGT GCAATCCATC TTGTTCAATC ATGCGAAACG ATCCTCATCC
5751 TGTCTCTTGA TCAGATCTTG ATCCCCTGCG CCATCAGATC
CTTGGCGGCA 5801 AGAAAGCCAT CCAGTTTACT TTGCAGGGCT
TCCCAACCTT ACCAGAGGGC 5851 GCCCCAGCTG GCAATTCCGG
TTCGCTTGCT GTCCATAAAA CCGCCCAGTC 5901 TAGCTATCGG
CATGTAAGCC CACTGCAAGC TACCTGCTTT CTCTTTGCGC 5951
TTGCGTTTTC CCTTGTCCAG ATAGCCCAGT AGCTGACATT CATCCGGGGT
6001 CAGCACCGTT TCTGCGGACT GGCTTTCTAC GTGTTCCGCT
TCCTTTAGCA 6051 GCCCTTGCGC CCTGAGTGCT TGCGGCAGCG
TGAAGCTTTT TGCAAAAGCC 6101 TAGGCCTCCA AAAAAGCCTC
CTCACTACTT CTGGAATAGC TCAGAGGCCG 6151 AGGCGGCCTC
GGCCTCTGCA TAAATAAAAA AAATTAGTCA GCCATGGGGC 6201
GGAGAATGGG CGGAACTGGG CGGAGTTAGG GGCGGGATGG GCGGAGTTAG
6251 GGGCGGGACT ATGGTTGCTG ACTAATTGAG ATGCATGCTT
TGCATACTTC 6301 TGCCTGCTGG GGAGCCTGGG GACTTTCCAC
ACCTGGTTGC TGACTAATTG 6351 AGATGCATGC TTTGCATACT
TCTGCCTGCT GGGGAGCCTG GGGACTTTCC 6401 ACACCCTAAC
TGACACACAT TCCACA

(93) As described herein, the vector may comprise two, three, four, five, or more nucleic acid sequences according to the present application. In some embodiments, the vector comprises a first nucleic acid sequences encoding a first RNA-regulated fusion protein and a second nucleic acid sequence encoding a second RNA-regulated fusion protein. In other embodiments, the vector may further comprise a third nucleic acid molecule encoding a third RNA-regulated fusion protein, etc. For example, the vector may comprise 3-10 or more nucleic acid molecules, each encoding an independently selected RNA fusion protein according to the present application.

(94) In some embodiments, where the vector encodes multiple RNA-regulated fusion proteins, each independent fusion protein may comprise a component of a metabolic pathway. In some embodiments, the metabolic pathway is glucose metabolism and the independent fusion proteins comprise insulin, glucagon, and/or protein kinase C epsilon. In other embodiments, the metabolic pathway is a GPCR signaling pathway and the independent fusion proteins are selected from the group consisting of α, β, and γ subunits of G-proteins.

(95) In other embodiments, where the vector encodes multiple RNA-regulated fusion proteins, each RNA-regulated fusion protein comprises a distinct protein of interest. Suitable proteins of interest are described in detail above. In some embodiments, the proteins of interest comprise fluorescent proteins. In accordance with such embodiments, the fluorescent proteins have fluorescent emission spectra that do not substantially overlap with one another.

(96) In some embodiments, the present application relates to an expression system comprising an expression vector into which is inserted a nucleic acid molecule described herein. In one embodiment, the expression system comprises a first vector encoding an RNA-regulated fusion protein and a second vector encoding a lentiviral transactivator of transcription (Tar) RNA aptamer.

(97) Some embodiments of the present application relate to a host cell comprising a nucleic acid molecule (i.e., a nucleic acid molecule encoding an RNA-regulated fusion protein and/or a lentiviral transactivator of transcription (Tar) RNA sequence) or a vector (i.e., a vector comprising a nucleic acid molecule encoding an RNA-regulated fusion protein and/or a lentiviral transactivator

of transcription (Tar) RNA sequence) described herein.
(98) In some embodiments, the host cell is a mammalian cell. Suitable mammalian cells include, without limitation, rodent cells (i.e., mouse or rat cells), rabbit cells, guinea pig cells, feline cells, canine cells, porcine cells, equine cells, bovine cell, ovine cells, monkey cells, non-human primate, or human cells. In some embodiments, the host cell is a human cell. Suitable cells comprising the nucleic acid molecule or vector as described herein include primary or immortalized embryonic cells, fetal cells, or adult cells, at any stage of their lineage, e.g., totipotent, pluripotent, multipotent, or differentiated cells.
(99) The nucleic acid molecules and/or vectors described herein may be introduced into cells via transformation, particularly transduction, conjugation, lipofection, protoplast fusion, mobilization, particle bombardment, microinjection, transfection, or electroporation. In some embodiments, the nucleic acid molecules described herein are incorporated into the host cell using standard cloning procedures known in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Springs Laboratory, Cold Springs Harbor, N.Y. (1989), which is hereby incorporated by reference in its entirety.
(100) In some embodiments, the host cell may comprise an endogenous RNA ligase. As described herein, the endogenous RNA ligase has the ability to catalyze the circularization of a ribonucleic acid molecule having a 5′-OH and a 2′-3′-cyclic phosphate. In accordance with this embodiment, the endogenous RNA ligase is RtcB.
(101) Another aspect of the present application relates to an RNA-regulated fusion protein comprising a protein of interest and an RNA-regulated destabilization domain. Suitable proteins of interest and RNA-regulated destabilization domains are described in more detail supra.
(102) In some embodiments, the protein of interest is a fluorescent protein, a bioluminescent protein, an enzyme, or a transcription factor. Suitable fluorescent proteins, bioluminescent proteins, enzymes, or transcription factors are described in more detail supra.
(103) In some embodiments, the RNA-regulated destabilization domain has the consensus sequence of SEQ ID NO: 62 as follows: XXXXXXXXXXXXXXXXXx, where X at position 1 can be S or A; X at position 2 can be G or A; X at position 3 can be P or A; X at position 4 can be R or K; X at position 5 can be P, A, I, Y, K, or R; X at position 6 can be R, K, V, or Y; X at position 7 can be G, A, or R; X at position 8 can be T or A; X at position 9 can be R or K; X at position 10 can be G or A; X at position 11 can be K or A; X at position 12 can be G or A; X at position 13 can be R or K; X at position 14 can be I or A; X at position 15 can be R, K, Y, or G; X at position 16 can be R, K, V, T, or Y; X at position 17 can be any amino acid but preferably R, G, E, S, or C; and x at position 18 is optional and can be any amino acid, but preferably G, E, O, N, D, or E.
(104) In some embodiments the RNA-regulated destabilization domain has the sequence of tDeg (SEQ ID NO: 63) as follows: SGPRPRGTRGKGRRIRRRG.
(105) Exemplary RNA-regulated fusion proteins are identified in Table 8 below.
(106) TABLE-US-00014 TABLE 8 Exemplary RNA-Regulated Fusion Proteins SEQ ID Vector Sequence NO: (mNeonGreen).sub.4-
MVSKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPNDGYE 80 tDeg
ELNLKSTKGDLQFSPWILVPHIGYGFHQYLPYPDGMSPFQAAMVD
GSGYQVHRTMQFEDGASLTVNYRYTYEGSHIKGEAQVKGTGFPAD
GPVMTNSLTAADWCRSKKTYPNDKTIISTFKWSYTTGNGKRYRST
ARTTYTFAKPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAF
TDVMGMDELYKGGHMGTGSTGGTGGVSKGEEDNMASLPATHELHI
FGSINGVDFDMVGQGTGNPNDGYEELNLKSTKGDLQFSPWILVPH
IGYGFHQYLPYPDGMSPFQAAMVDGSGYQVHRTMQFEDGASLTVN
YRYTYEGSHIKGEAQVKGTGFPADGPVMTNSLTAADWCRSKKTYP
NDKTIISTFKWSYTTGNGKRYRSTARTTYTFAKPMAANYLKNQPM
YVFRKTELKHSKTELNFKEWQKAFTDVMGMDELYKSGLESSGGTG

GSGGVSKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPND
GYEELNLKSTKGDLQFSPWILVPHIGYGFHQYLPYPDGMSPFQAA
MVDGSGYQVHRTMQFEDGASLTVNYRYTYEGSHIKGEAQVKGTGF
PADGPVMTNSLTAADWCRSKKTYPNDKTIISTFKWSYTTGNGKRY
RSTARTTYTFAKPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQ
KAFTDVMGMDELYKGGSGTGGTASSGSGGGVSKGEEDNMASLPAT
HELHIFGSINGVDFDMVGQGTGNPNDGYEELNLKSTKGDLQFSPW
ILVPHIGYGFHQYLPYPDGMSPFQAAMVDGSGYQVHRTMQFEDGA
SLTVNYRYTYEGSHIKGEAQVKGTGFPADGPVMTNSLTAADWCRS
KKTYPNDKTIISTFKWSYTTGNGKRYRSTARTTYTFAKPMAANYL
KNQPMYVFRKTELKHSKTELNFKEWQKAFTDVMGMDELYKGGRSG
GGSGPRPRGTRGKGRRIRRRG (GenBank Accession No. QEM23463.1 and GenBank Accession No. QEM23465.1, which are hereby incorporated by reference in their entirety) mNeonGreen-
MVSKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPNDGYE 81 tDeg
ELNLKSTKGDLQFSPWILVPHIGYGFHQYLPYPDGMSPFQAAMVD
GSGYQVHRTMQFEDGASLTVNYRYTYEGSHIKGEAQVKGTGFPAD
GPVMTNSLTAADWCRSKKTYPNDKTIISTFKWSYTTGNGKRYRST
ARTTYTFAKPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAF
TDVMGMDELYKGGHMGGGSGGGSGPRPRGTRGKGRRIRRRG mCherry-tDeg
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEG 82
TQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKL
SFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDA
EVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYKGGSGGGSGPRPRGTRGKGRRIRRRG NanoLuc-tDeg
MVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQR 83
IVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVYPVDDHHF
KVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGILWN
GNKIIDERLINPDGSLLFRVTINGVTGWRLCERILAGGSHMGGSG
GGSGPRPRGTRGKGRRIRRRG EYFP-tDeg
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 84
KFICTIGKLPVPWPTLVITFGYGLQCFARYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGGSGGGSGPRPRGTRGKGRRIRRRG EGFP-TetR-
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 85 tDeg
KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGTGACGTSGGRLDKSKVINSALELLNEVGIE
GLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEMLDRHHTHFC
PLEGESWQDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETL
ENQLAFLCQQGFSLENALYALSAVGHFTLGCVLEDQEHQVAKEER
ETPTTDSMPPLLRQAIELFDHQGAEPAFLFGLELIICGLEKQLKC
ESGSGSGTGGIGGSGPRPRGTRGKGRRIRRRG mCherry-TetR-
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEG 86 tDeg
TQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKL

SFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDA
EVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYKGTGACGTSGGRLDKSKVINSALELLNEVGIEGLT
TRKLAQKLGVEQPTLYWHVKNKRALLDALAIEMLDRHHTHFCPLE
GESWQDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQ
LAFLCQQGFSLENALYALSAVGHFTLGCVLEDQEHQVAKEERETP
TTDSMPPLLRQAIELFDHQGAEPAFLFGLELIICGLEKQLKCESG
SGSGTGGIGGSGPRPRGTRGKGRRIRRRG EGFP-EZH2-
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 87 tDeg
KFICTIGKLPVPWPTLVTILTYGVQCFSRYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGTGACGTSGGMGQTGKKSEKGPVCWRKRVKS
EYMRLRQLKRFRRADEVKSMFSSNRQKILERTEILNQEWKQRRIQ
PVHILTSVSSLRGTRECSVTSDLDFPTQVIPLKTLNAVASVPIMY
SWSPLQQNFMVEDETVLHNIPYMGDEVLDQDGTFIEELIKNYDGK
VHGDRECGFINDEIFVELVNALGQYNDDDDDDDGDDPEEREEKQK
DLEDHRDDKESRPPRKFPSDKIFEAISSMFPDKGTAEELKEKYKE
LTEQQLPGALPPECTPNIDGPNAKSVQREQSLHSFHTLFCRRCFK
YDCFLHPFHATPNTYKRKNTETALDNKPCGPQCYQHLEGAKEFAA
ALTAERIKTPPKRPGGRRRGRLPNNSSRPSTPTINVLESKDTDSD
REAGTETGGENNDKEEEEKKDETSSSSEANSRCQTPIKMKPNIEP
PENVEWSGAEASMFRVLIGTYYDNFCAIARLIGTKTCRQVYEFRV
KESSIIAPAPAEDVDTPPRKKKRKHRLWAAHCRKIQLKKDGSSNH
VYNYQPCDHPRQPCDSSCPCVIAQNFCEKFCQCSSECQNRFPGCR
CKAQCNTKQCPCYLAVRECDPDLCLTCGAADHWDSKNVSCKNCSI
QRGSKKHLLLAPSDVAGWGIFIKDPVQKNEFISEYCGEIISQDEA
DRRGKVYDKYMCSFLFNLNNDFVVDATRKGNKIRFANHSVNPNCY
AKVMMVNGDHRIGIFAKRAIQTGEELFFDYRYSQADALKYVGIER
EMEIPGSGTGGIGGSGPRPRGTRGKGRRIRRRG mCherry-
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEG 88 EZH2-tDeg
TQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKL
SFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDA
EVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYKGTGACGTSGGMGQTGKKSEKGPVCWRKRVKSEYM
RLRQLKRFRRADEVKSMFSSNRQKILERTEILNQEWKQRRIQPVH
ILTSVSSLRGTRECSVTSDLDFPTQVIPLKTLNAVASVPIMYSWS
PLQQNFMVEDETVLHNIPYMGDEVLDQDGTFIEELIKNYDGKVHG
DRECGFINDEIFVELVNALGQYNDDDDDDDGDDPEEREEKQKDLE
DHRDDKESRPPRKFPSDKIFEAISSMFPDKGTAEELKEKYKELTE
QQLPGALPPECTPNIDGPNAKSVQREQSLHSFHTLFCRRCFKYDC
FLHPFHATPNTYKRKNTETALDNKPCGPQCYQHLEGAKEFAAALT
AERIKTPPKRPGGRRRGRLPNNSSRPSTPTINVLESKDTDSDREA
GTETGGENNDKEEEEKKDETSSSSEANSRCQTPIKMKPNIEPPEN
VEWSGAEASMFRVLIGTYYDNFCAIARLIGTKTCRQVYEFRVKES
SIIAPAPAEDVDTPPRKKKRKHRLWAAHCRKIQLKKDGSSNHVYN
YQPCDHPRQPCDSSCPCVIAQNFCEKFCQCSSECQNRFPGCRCKA

QCNTKQCPCYLAVRECDPDLCLTCGAADHWDSKNVSCKNCSIQRG
SKKHLLLAPSDVAGWGIFIKDPVQKNEFISEYCGEIISQDEADRR
GKVYDKYMCSFLFNLNNDFVVDATRKGNKIRFANHSVNPNCYAKV
MMVNGDHRIGIFAKRAIQTGEELFFDYRYSQADALKYVGIEREME
IPGSGTGGTGGSGPRPRGTRGKGRRIRRRG EGFP-NFκB-
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 89 tDeg
KFICTIGKLPVPWPTLVTILTYGVQCFSRYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGGSGGSGGSGGSGGTGAEDDPYLGRPEQMFH
LDPSLTHTIFNPEVFQPQMALPTADGPYLQILEQPKQRGFRFRYV
CEGPSHGGLPGASSEKNKKSYPQVKICNYVGPAKVIVQLVTNGKN
IHLHAHSLVGKHCEDGICTVTAGPKDMVVGFANLGILHVTKKKVF
ETLEARMTEACIRGYNPGLLVHPDLAYLQAEGGGDRQLGDREKEL
IRQAALQQTKEMDLSVVRLMFTAFLPDSTGSFTRRLEPVVSDAIY
DSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIRFY
EEEENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDINITKPASVF
VQLRRKSDLETSEPKPFLYYPEIKDKEEVQRKRQKLMPNFSDSFG
GGSGAGAGGGGMFGSGGGGGGTGSTGPGYSFPHYGFPTYGGITFH
PGTTKSNAGMKHGTMDTESKKDPEGCDKSDDKNTVNLFGKDPRGS
LSGGTGGSGPRPRGTRGKGRRIRRRG mCherry-
MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEG 90 NFκB-tDeg
TQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKL
SFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDA
EVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYKGGSGGSGGSGGSGGTGAEDDPYLGRPEQMFHLDP
SLTHTIFNPEVFQPQMALPTADGPYLQILEQPKQRGFRFRYVCEG
PSHGGLPGASSEKNKKSYPQVKICNYVGPAKVIVQLVTNGKNIHL
HAHSLVGKHCEDGICTVTAGPKDMVVGFANLGILHVTKKKVFETL
EARMTEACIRGYNPGLLVHPDLAYLQAEGGGDRQLGDREKELIRQ
AALQQTKEMDLSVVRLMFTAFLPDSTGSFTRRLEPVVSDAIYDSK
APNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIRFYEEE
ENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDINITKPASVFVQL
RRKSDLETSEPKPFLYYPEIKDKEEVQRKRQKLMPNFSDSFGGGS
GAGAGGGGMFGSGGGGGGTGSTGPGYSFPHYGFPTYGGITFHPGT
TKSNAGMKHGTMDTESKKDPEGCDKSDDKNTVNLFGKDPRGSLSG
GTGGSGPRPRGTRGKGRRIRRRG EGFP-
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 91 TurboID-tDeg
KFICTIGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGTGACGTSGGMKDNTVPLKLIALLANGEFHS
GEQLGETLGMSRAAINKHIQTLRDWGVDVFTVPGKGYSLPEPIPL
LNAKQILGQLDGGSVAVLPVVDSTNQYLLDRIGELKSGDACIAEY
QQAGRGSRGRKWFSPFGANLYLSMFWRLKRGPAAIGLGPVIGIVM
AEALRKLGADKVRVKWPNDLYLQDRKLAGILVELAGITGDAAQIV
IGAGINVAMRRVEESVVNQGWITLQEAGINLDRNTLAATLIRELR

AALELFEQEGLAPYLPRWEKLDNFINRPVKLIIGDKEIFGISRGI
DKQGALLLEQDGVIKPWMGGEISLRSAEKGSGTGGTGGSGPRPRG TRGKGRRIRRRG
EGFP-APEX- MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL 92 tDeg
KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMP
EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG
NILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA
DHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKGTGACGTSGKSYPTVSADYQDAVEKAKKKLR
GFIAEKRCAPLMLRLAFHSAGTFDKGTKTGGPFGTIKHPAELAHS
ANNGLDIAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPKVP
FHPGREDKPEPPPEGRLPDPTKGSDHLRDVFGKAMGLTDQDIVAL
SGGHTIGAAHKERSGFEGPWTSNPLIFDNSYFTELLSGEKEGLLQ
LPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFAD
AGSGTGGTGGSGPRPRGTRGKGRRIRRRG

(107) Yet another aspect of the disclosure relates to a molecular complex comprising an RNA-regulated fusion protein comprising (i) a protein of interest and (ii) an RNA-regulated destabilization domain and an RNA aptamer bound specifically to the RNA-regulated destabilization domain.

(108) In some embodiments, the protein of interest is a fluorescent protein, a bioluminescent protein, an enzyme, or a transcription factor. Suitable fluorescent proteins, bioluminescent proteins, enzymes, and transcription factors are described in detail supra.

(109) In some embodiments, the RNA-regulated destabilization domain has the sequence of SEQ ID NO: 62, where X at position 1 is S or A; X at position 2 is G or A; X at position 3 is P or A; X at position 4 is R or K; X at position 5 is P, A, I, Y, K, or R; X at position 6 is R, K, V, or Y; X at position 7 is G, A, or R; X at position 8 is T or A; X at position 9 is R or K; X at position 10 is G or A; X at position 11 is K or A; X at position 12 is G or A; X at position 13 is R or K; X at position 14 is I or A; X at position 15 is R, K, Y, or G; X at position 16 is R, K, V, T, or Y; X at position 17 is any amino acid; and x at position 18 is optional and can be any amino acid. For example, the RNA-regulated destabilization domain may be tDeg (SEQ ID NO: 63).

(110) Suitable RNA aptamer sequences are described in detail supra. In some embodiments, the RNA aptamer comprises the consensus sequence of SEQ ID NO: 56, SEQ ID NO: 58, or SEQ ID NO: 60, wherein N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. For example, the RNA aptamer may comprise the sequence of wild-type TAR RNA (SEQ ID NO: 57), TAR Variant-1 (SEQ ID NO: 59), or TAR Variant-2 (Pepper; SEQ ID NO: 61).

(111) Additional exemplary RNA aptamers may be selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, and SEQ ID NO: 73.

(112) Some embodiments of the present application relate to a host cell comprising a molecular complex described herein (i.e., a molecular complex comprising an RNA-regulated fusion protein and an RNA aptamer bound specifically to the RNA-regulated destabilization domain). Suitable host cells are described in detail supra.

(113) In some embodiments, the host cell is a mammalian cell. As described herein above, suitable mammalian cells include, without limitation, rodent cells (i.e., mouse or rat cells), rabbit cells, guinea pig cells, feline cells, canine cells, porcine cells, equine cells, bovine cell, ovine cells, monkey cells, non-human primate, or human cells. In some embodiments, the host cell is a human cell.

(114) Another aspect of the invention relates to a method of imaging RNA in a cell. This method involves providing a first vector encoding an RNA-regulated fusion protein, wherein the RNA-regulated fusion protein comprises a fluorescent protein, a bioluminescent protein, or an enzyme fused to an RNA-regulated destabilization domain; providing second vector encoding an RNA

molecule comprising (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; transfecting a host cell with the first vector and the second vector; and imaging said contacted cells.
(115) Suitable vectors for carrying out the methods of imaging RNA in a cell are described in more detail supra and include, e.g., a plasmid (e.g., an expression vector) and a viral vector (e.g., a lentiviral or adenoviral vector).
(116) Suitable RNA-regulated fusion proteins for carrying out the methods of the present application are described in more detail supra. In some embodiments of the methods described herein, the RNA-regulated fusion protein is a fluorescent protein selected from the group consisting of Green Fluorescent Protein, Enhanced Green Fluorescent Protein (EGFP), Enhanced Yellow Fluorescent Protein (EYFP), Venus, mVenus, Citrine, mCitrine, Cerulean, mCerulean, Orange Fluorescent Protein (OFP), mNeonGreen, moxNeonGreen, mCherry, mTagBFP, Venus, mVenus, mTurquoise, mScarlet, mWasabi, mOrange, and dTomato.
(117) In other embodiments of the methods described herein, the RNA-regulated fusion protein is a bioluminescent protein selected from the group consisting of luciferase, β-galactosidase, β-lactamase, peroxidase, alkaline phosphatase, β-glucuronidase, and β-glucosidase. In some embodiments, the bioluminescent protein is a luciferase selected from the group consisting of Nanoluc luciferase (Nluc), Firefly luciferase, and *Renilla* luciferase (Rluc).
(118) In further embodiments of the methods described herein, the RNA-regulated fusion protein is an enzyme, wherein the enzyme is a biotin ligase. Suitable biotin ligases are described in detail supra and include, e.g., TurboID, miniTurbo, or *E. coli* BirA.
(119) As described in more detail supra, the RNA-regulated destabilization domain may comprise a bifunctional peptide having a lentiviral transactivator of transcription (Tat) peptide and a degron peptide. Lentiviral transactivator of transcription (Tat) peptides and a degron peptides are described in more detail supra.
(120) In some embodiments of the methods described herein, the RNA-regulated destabilization domain comprises the consensus sequence of SEQ ID NO: 62, where X at position 1 is S or A; X at position 2 is G or A; X at position 3 is P or A; X at position 4 is R or K; X at position 5 is P, A, I, Y, K, or R; X at position 6 is R, K, V, or Y; X at position 7 is G, A, or R; X at position 8 is T or A; X at position 9 is R or K; X at position 10 is G or A; X at position 11 is K or A; X at position 12 is G or A; X at position 13 is R or K; X at position 14 is I or A; X at position 15 is R, K, Y, or G; X at position 16 is R, K, V, T, or Y; X at position 17 is any amino acid; and x at position 18 is optional and can be any amino acid. Thus, in some embodiments, the RNA-regulated destabilization domain is tDeg (SEQ ID NO: 63).
(121) As used herein, an RNA of interest is an RNA molecule that is desired and/or is being assessed. The RNA of interest may be a messenger RNA (mRNA) or a noncoding RNA (ncRNA). A messenger RNA or “mRNA” refers to a single-stranded RNA molecule that specifies the amino acid sequence of a protein. The mRNA molecule may comprise a 5′ untranslated region (5′ UTR), a coding region, and a 3′ untranslated region (3′ UTR). A 5′ UTR is an untranslated nucleotide segment in an RNA molecule immediately preceding the AUG start codon. A 3′ UTR is an untranslated nucleotide segment in an RNA molecule immediately following the translation termination codon.
(122) In some embodiments, the RNA of interest is an mRNA and the RNA aptamer is located within a coding region of the mRNA. In other embodiment, the RNA of interest is a mRNA and the RNA aptamer is located upstream of the 5′ UTR, within the 5′ UTR, within the 3′ UTR, or downstream of the 3′ UTR.
(123) In other embodiments, the RNA of interest is a noncoding RNA (ncRNA). As described herein, a noncoding RNA refers to a functional RNA molecule that is not translated into a protein. The RNA of interest may be a noncoding RNA selected from the group consisting of ribosomal RNA (rRNA), transfer RNA (tRNA), heterogeneous nuclear RNA (hnRNA), small cytoplasmic

RNA (scRNA), small nuclear (snRNA), small nucleolar (snoRNA), ribozymes, and regulatory RNA (e.g., siRNA, miRNA, microRNA, etc.).
(124) In some embodiments, the RNA of interest is an artificial, engineered synthetic RNA.
(125) Suitable RNA aptamers are described in detail supra. In some embodiments of the methods described herein, the RNA aptamer comprises the consensus sequence of SEQ ID NO: 56, SEQ ID NO: 58, or SEQ ID NO: 60, where N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. For example, the RNA aptamer may comprise the sequence of wild-type TAR RNA (SEQ ID NO: 57), TAR Variant-1 (SEQ ID NO: 59), or TAR Variant-2 (Pepper; SEQ ID NO: 61). In some embodiments of the methods described herein, the RNA aptamer comprises the sequence of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, or SEQ ID NO: 73.
(126) Methods of transfecting a host cell are well known in the art and described in more detail supra. According to some embodiments of the methods described herein, transfecting the host cell with the first vector and the second vector is carried out simultaneously. In other embodiments, transfecting the host cell with the first vector and the second vector is carried out sequentially.
(127) Methods of imaging cells are well known in the art. In some embodiments, imaging said transfected cells is carried out by fluorescence microscopy or imaging flow cytometry (see, e.g., Wu et al., "Live Imaging of mRNA Using RNA-Stabilized Fluorogenic Proteins," *Nature Methods* 16:862-565 (2019) and Wu & Jaffrey, Live Imaging of mRNA Using Pepper RNA-Stabilized Fluorgenic Proteins," *Nature Methods*, DOI: 10.21203/rs.2.11494/v1 (2019), which are hereby incorporated by reference in their entirety).
(128) Yet another aspect of the invention relates to a method of imaging RNA in a cell. This method involves providing a vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a fluorescent protein, a bioluminescent protein, or an enzyme fused to an RNA-regulated destabilization domain; transfecting a host cell with the first vector; contacting said transfected cell with an RNA molecule comprising (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; and imaging said contacted cells.
(129) Suitable vectors for carrying out the methods of imaging RNA in a cell are described in more detail supra and include, e.g., a plasmid (e.g., an expression vector) and a viral vector (e.g., a lentiviral or adenoviral vector).
(130) Suitable RNA-regulated fusion proteins for carrying out the methods of the present application are described in more detail supra. In some embodiments of the methods described herein, the RNA-regulated fusion protein is a fluorescent protein selected from the group consisting of Green Fluorescent Protein, Enhanced Green Fluorescent Protein (EGFP), Enhanced Yellow Fluorescent Protein (EYFP), Venus, mVenus, Citrine, mCitrine, Cerulean, mCerulean, Orange Fluorescent Protein (OFP), mNeonGreen, moxNeonGreen, mCherry, mTagBFP, Venus, mVenus, mTurquoise, mScarlet, mWasabi, mOrange, and dTomato.
(131) In other embodiments of the methods described herein, the RNA-regulated fusion protein is a bioluminescent protein selected from the group consisting of luciferase, β-galactosidase, β-lactamase, peroxidase, alkaline phosphatase, β-glucuronidase, and β-glucosidase. In some embodiments, the bioluminescent protein is a luciferase selected from the group consisting of Nanoluc luciferase (Nluc), Firefly luciferase, and *Renilla* luciferase (Rluc).
(132) In further embodiments of the methods described herein, the RNA-regulated fusion protein is an enzyme, wherein the enzyme is a biotin ligase. Suitable biotin ligases are described in detail supra and include, e.g., TurboID, miniTurbo, or *E. coli* BirA.
(133) As described in more detail supra, the RNA-regulated destabilization domain may comprise a bifunctional peptide having a lentiviral transactivator of transcription (Tat) peptide and a degron peptide. Lentiviral transactivator of transcription (Tat) peptides and a degron peptides are described in more detail supra.

(134) In some embodiments of the methods described herein, the RNA-regulated destabilization domain comprises the consensus sequence of SEQ ID NO: 62, where X at position 1 is S or A; X at position 2 is G or A; X at position 3 is P or A; X at position 4 is R or K; X at position 5 is P, A, I, Y, K, or R; X at position 6 is R, K, V, or Y; X at position 7 is G, A, or R; X at position 8 is T or A; X at position 9 is R or K; X at position 10 is G or A; X at position 11 is K or A; X at position 12 is G or A; X at position 13 is R or K; X at position 14 is I or A; X at position 15 is R, K, Y, or G; X at position 16 is R, K, V, T, or Y; X at position 17 is any amino acid; and x at position 18 is optional and can be any amino acid. Thus, in some embodiments, the RNA-regulated destabilization domain is tDeg (SEQ ID NO: 63).
(135) In some embodiments, the RNA of interest is a mRNA and the RNA aptamer is located within a coding region of the mRNA. In other embodiment, the RNA of interest is a mRNA and the RNA aptamer is located upstream of the 5′ UTR, within the 5′ UTR, within the 3′ UTR, or downstream of the 3′ UTR.
(136) In other embodiments, the RNA of interest is a noncoding RNA (ncRNA). As described herein, the term “noncoding RNA” refers to a functional RNA molecule that is not translated into a protein. The RNA of interest may be a noncoding RNA selected from the group consisting of ribosomal RNA (rRNA), transfer RNA (tRNA), heterogeneous nuclear RNA (hnRNA), small cytoplasmic RNA (scRNA), small nuclear (snRNA), small nucleolar (snoRNA), ribozymes, and regulatory RNA (e.g., siRNA, miRNA, microRNA, etc.).
(137) Suitable RNA aptamers are described in detail supra. In some embodiments of the methods described herein, the RNA aptamer comprises the consensus sequence of SEQ ID NO: 56, SEQ ID NO: 58, or SEQ ID NO: 60, wherein N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. For example, the RNA aptamer may comprise the sequence of wild-type TAR RNA (SEQ ID NO: 57), TAR Variant-1 (SEQ ID NO: 59), or TAR Variant-2 (Pepper; SEQ ID NO: 61). In some embodiments of the methods described herein, the RNA aptamer comprises the sequence of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, or SEQ ID NO: 73.
(138) The RNA molecule comprising the (i) RNA sequence of interest and (ii) the RNA aptamer sequence may be a circular RNA molecule or a linear RNA molecule.
(139) Methods of transfecting a host cell are well known in the art and described in more detail supra.
(140) Contacting the transfected cell may be carried out by allowing the RNA molecule comprising the (i) RNA sequence of interest and (ii) the RNA aptamer sequence may be a circular RNA molecule or a linear RNA molecule to diffuse into the cell.
(141) Methods of imaging cells are well known in the art. In some embodiments, imaging said contacted cells is carried out by fluorescence microscopy or imaging flow cytometry.
(142) A further aspect of the invention relates to a method of selectively modifying an RNA-binding protein. This method involves providing a first expression vector encoding a RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises an enzyme fused to an RNA-regulated destabilization domain; providing a second expression vector encoding (i) an RNA sequence of interest and (ii) an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequences; transfecting a host cell with the first and second expression vectors; and allowing the enzyme to be expressed, wherein the expressed enzyme selectively modifies a protein that binds to the RNA sequence of interest.
(143) Suitable enzymes are described in more detail supra. In some embodiments, the enzyme is selected from the group consisting of a ligase, a peroxidase, and a methyltransferase.
(144) In some embodiments of the methods described herein, the enzyme is a biotin ligase selected from the group consisting of TurboID, miniTurbo, and *E. coli* BirA.
(145) In some embodiments of the methods described herein, the enzyme is a peroxidase selected from the group consisting of an ascorbate peroxidase and a horseradish peroxidase. The ascorbate

peroxidase may be APEX2.
(146) As described in more detail supra, the RNA-regulated destabilization domain may comprise a bifunctional peptide having a lentiviral transactivator of transcription (Tat) peptide and a degron peptide. Lentiviral transactivator of transcription (Tat) peptides and a degron peptides are described in more detail supra.
(147) In some embodiments of the methods described herein, the RNA-regulated destabilization domain comprises the consensus sequence of SEQ ID NO: 62, where X at position 1 is S or A; X at position 2 is G or A; X at position 3 is P or A; X at position 4 is R or K; X at position 5 is P, A, I, Y, K, or R; X at position 6 is R, K, V, or Y; X at position 7 is G, A, or R; X at position 8 is T or A; X at position 9 is R or K; X at position 10 is G or A; X at position 11 is K or A; X at position 12 is G or A; X at position 13 is R or K; X at position 14 is I or A; X at position 15 is R, K, Y, or G; X at position 16 is R, K, V, T, or Y; X at position 17 is any amino acid; and x at position 18 is optional and can be any amino acid. Thus, in some embodiments, the RNA-regulated destabilization domain is tDeg (SEQ ID NO: 63).
(148) In some embodiments, the RNA of interest is a mRNA and the RNA aptamer is located within a coding region of the mRNA. In other embodiment, the RNA of interest is a mRNA and the RNA aptamer is located upstream of the 5′ UTR, within the 5′ UTR, within the 3′ UTR, or downstream of the 3′ UTR.
(149) In other embodiments, the RNA of interest is a noncoding RNA (ncRNA). As described herein, the term “noncoding RNA” refers to a functional RNA molecule that is not translated into a protein. The RNA of interest may be a noncoding RNA selected from the group consisting of ribosomal RNA (rRNA), transfer RNA (tRNA), heterogeneous nuclear RNA (hnRNA), small cytoplasmic RNA (scRNA), small nuclear (snRNA), small nucleolar (snoRNA), ribozymes, and regulatory RNA (e.g., siRNA, miRNA, microRNA, etc.).
(150) Suitable RNA aptamers are described in detail supra. In some embodiments of the methods described herein, the RNA aptamer comprises the consensus sequence of SEQ ID NO: 56, SEQ ID NO: 58, or SEQ ID NO: 60, wherein N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. For example, the RNA aptamer may comprise the sequence of wild-type TAR RNA (SEQ ID NO: 57), TAR Variant-1 (SEQ ID NO: 59), or TAR Variant-2 (Pepper; SEQ ID NO: 61). In some embodiments of the methods described herein, the RNA aptamer comprises the sequence of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, or SEQ ID NO: 73.
(151) In some embodiments of the methods of selectively modifying an RNA-binding protein described herein, the method further involves identifying a protein that is selectively modified by the enzyme within the transfected cells. See, e.g., Ramanathan et al., “RNA-Protein Interaction Detection in Living Cells,” *Nature Methods* 15:207-212 (2018), which is hereby incorporated by reference in its entirety.
(152) Another aspect relates to a method of regulating expression of an RNA-stabilized protein of interest. This method involves providing a first vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a protein of interest fused to an RNA-regulated destabilization domain; providing a second vector encoding an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; providing a host cell comprising a functional ubiquitination system; transfecting the host cell with the first and second expression vectors; and expressing the first and second expression vectors within the host cell, where said expressing the first and second expression vectors regulates proteomic stability of the RNA-regulated fusion protein; and where, in the absence of any expressed RNA aptamer sequence in the host cell, the RNA-regulated destabilization domain promotes degradation of the RNA-regulated fusion protein by the ubiquitination system; and where the RNA-regulated fusion protein is stabilized by the expressed RNA aptamer sequence.
(153) Another aspect of the invention relates to a method of regulating expression of an RNA-

stabilized protein of interest. This method involves providing a first vector encoding an RNA-regulated fusion protein, where the RNA-regulated fusion protein comprises a protein of interest fused to an RNA-regulated destabilization domain; providing a second vector encoding an RNA aptamer sequence, where the RNA-regulated destabilization domain specifically binds to the RNA aptamer sequence; providing a mammalian cell lysate or solution comprising (i) a ubiquitin ligase, (ii) proteosomal degradation machinery, (iii) transcriptional machinery, and (iv) translational machinery; contacting the mammalian cell lysate or solution with the first and second expression vectors; and expressing the first and second expression vectors, where said expressing the first and second expression vectors regulates proteomic stability of the RNA-regulated fusion protein; and where, in the absence of any expressed RNA aptamer sequence in the cell lysate or solution, the RNA-regulated destabilization domain promotes degradation of the RNA-regulated fusion protein by the proteosomal degradation system; and where the RNA-regulated fusion protein is stabilized by the expressed RNA aptamer sequence.

(154) Suitable proteins of interest for use in the methods described herein are described in more detail supra. In some embodiments, the protein of interest is a fluorescent protein, a bioluminescent protein, an enzyme, or a transcription factor. In other embodiments, the protein of interest is selected from the group consisting of a G-protein coupled receptor (GPCR), a nuclear receptor, a voltage gated ion channel, a ligand gated channel, a receptor tyrosine kinase, a growth factor, a phosphatase, a protein kinase, a viral regulator, a bacterial cell division protein, a scaffold protein, a DNA repair protein, a cytoskeletal protein, a ribosome, a histone deacetylase, an apoptosis regulator, a chaperone protein, a kinase, a phosphorylase, a phosphatase, deacetylase, a cytoskeletal protein (e.g., myosin, actin, dynein, kinesin, and tubulin).

(155) Suitable expression vectors encoding RNA-regulated fusion proteins and vectors encoding an RNA aptamer sequence for use in the methods described herein are described in detail supra and include, e.g., a plasmid (e.g., an expression vector) and a viral vector (e.g., a lentiviral or adenoviral vector).

(156) As described in more detail supra, the RNA-regulated destabilization domain may comprise a bifunctional peptide having a lentiviral transactivator of transcription (Tat) peptide and a degron peptide. Lentiviral transactivator of transcription (Tat) peptides and a degron peptides are described in more detail supra.

(157) In some embodiments of the methods described herein, the RNA-regulated destabilization domain comprises the consensus sequence of SEQ ID NO: 62, where X at position 1 is S or A; X at position 2 is G or A; X at position 3 is P or A; X at position 4 is R or K; X at position 5 is P, A, I, Y, K, or R; X at position 6 is R, K, V, or Y; X at position 7 is G, A, or R; X at position 8 is T or A; X at position 9 is R or K; X at position 10 is G or A; X at position 11 is K or A; X at position 12 is G or A; X at position 13 is R or K; X at position 14 is I or A; X at position 15 is R, K, Y, or G; X at position 16 is R, K, V, T, or Y; X at position 17 is any amino acid; and x at position 18 is optional and can be any amino acid. Thus, in some embodiments, the RNA-regulated destabilization domain is tDeg (SEQ ID NO: 63).

(158) Suitable RNA aptamer sequences for use in the methods described herein are described in more detail supra. In some embodiments, the RNA aptamer comprises the consensus sequence of SEQ ID NO: 56, SEQ ID NO: 58, or SEQ ID NO: 60, wherein N can be A, C, G, or U; S can be C or G; H can be A, C, or U; Y can be C or U; W can be A or U; B can be C, G, or U; M can be A or C; and D can be A, G, or U. For example, the RNA aptamer may comprises the sequence of wild-type TAR RNA (SEQ ID NO: 57), TAR Variant-1 (SEQ ID NO: 59), or TAR Variant-2 (Pepper; SEQ ID NO: 61). In other embodiments, the RNA aptamer comprises the sequence of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, or SEQ ID NO: 73.

(159) Suitable host cells for use in the methods described herein are described in more detail supra. In some embodiments, the host cell is a mammalian cell.

(160) Suitable mammalian cell lysates include, for example and without limitation, human cell

lysates, non-human primate cell lysates, feline cell lysates, canine cell lysates, ovine cell lysates, hircine cell lysates, bovine cell lysates, equine cell lysates, porcine cell lysates, leporine cell lysates, and murine cell lysates.

(161) Suitable solutions comprising (i) a ubiquitin ligase, (ii) proteosomal degradation machinery, (iii) transcriptional machinery, and (iv) translational machinery are well known in the art.

(162) Exemplary ubiquitin ligases include, without limitation, ubiquitin E3 ligases (Li et al., "Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, A Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling," *PLoS One* 3(1):e1487 (2008); Berndsen & Wolberger, "New Insights into Ubiquitin E3 Ligase Mechanism," *Nat. Struct. Mol. Biol.* 21(4):301-307 (2014), which are hereby incorporated by reference in their entirety). In some embodiments, the ubiquitin E3 ligase is selected form the group consisting of Really Interesting New Gene/U-box (RING) E3 ligase, Homologous to E6AP C-Terminus (HECT) E3 ligase, and RING between RING (RBR) E3 ligase (see, e.g., Metzger et al., "RING-Type E3 Ligases: Master Manipulators of E2 Ubiquitin-Conjugating Enzymes and Ubiquitination," *Biochim. Biophys. Acta.* 1843(1):47-60 (2014); Rotin & Kumar, "Physiological Functions of the HECT Family of Ubiquitin Ligases," *Nat. Rev. Mol. Cell. Biol.* 10(6):398-409 (2009); Sluimer & Distel, "Regulating the Human HECT E3 Ligases," *Cell Mol. Life Sci.* 75(17):3121-3141 (2018); Reiter & Klevit, "Characterization of RING-Between-RING E3 Ubiquitin Transfer Mechanisms," *Methods. Mol. Biol.* 1844:3-17 (2018); and Dove & Klevit, "RING-Between-RING E3 Ligases: Emerging Themes Amid the Variations," *J. Mol. Biol.* 429(22):3363-3375 (2017), which are hereby incorporated by reference in their entirety).

(163) Methods of transfecting cells are well known in the art and described in more detail supra.

(164) Another aspect of the present application relates to a treatment method. This method involves contacting a cell with an RNA aptamer, where upon said contacting, the aptamer interacts with an RNA-regulated destabilization domain fused to a protein of interest in the cell to stabilize the protein of interest in the cell.

(165) According to one embodiment, this and other treatment methods described herein are effective to treat a cell, e.g., a cell under a stress or disease condition. Exemplary cell stress conditions may include, without limitation, exposure to a toxin; exposure to chemotherapeutic agents, irradiation, or environmental genotoxic agents such as polycyclic hydrocarbons or ultraviolet (UV) light; exposure of cells to conditions such as glucose starvation, inhibition of protein glycosylation, disturbance of Ca2+ homeostasis and oxygen; exposure to elevated temperatures, oxidative stress, or heavy metals; and exposures to a pathological disease state (e.g., diabetes, Parkinson's disease, cardiovascular disease (e.g., myocardial infarction, end-stage heart failure, arrhythmogenic right ventricular dysplasia, and Adriamycin-induced cardiomyopathy), and various cancers (Fulda et al., "Cellular Stress Responses: Cell Survival and Cell Death," *Int. J Cell Biol*. (2010), which is hereby incorporated by reference in its entirety).

(166) In some embodiments, contacting a cell with an RNA molecule (aptamer) of the present application involves introducing an RNA molecule into a cell. Suitable methods of introducing RNA molecules into cells are well known in the art and include, but are not limited to, the use of transfection reagents, electroporation, microinjection, or via viruses.

(167) The cell may be a eukaryotic cell. Exemplary eukaryotic cells include a yeast cell, an insect cell, a fungal cell, a plant cell, and an animal cell (e.g., a mammalian cell). Suitable mammalian cells include, for example without limitation, human, non-human primate, cat, dog, sheep, goat, cow, horse, pig, rabbit, and rodent cells.

(168) The RNA molecule of the present invention may be isolated or present in in vitro conditions for extracellular expression and/or processing. According to this embodiment, the RNA molecule is contacted by an RNAligase (e.g., RtcB) in vitro, purified, circularized, and then the circularized RNA molecule is administered to a cell or subject for treatment.

(169) Treating cells also includes treating the organism in which the cells reside. Thus, by this and

the other treatment methods of the present invention, it is contemplated that treatment of a cell includes treatment of a subject in which the cell resides.
(170) In some embodiments, the treatment method further comprises introducing the protein of interest into the cell prior to said contacting.
(171) In some embodiments, the cell is in a patient.
(172) In some embodiments, introducing is carried out by any one or more of injecting mRNA encoding for the protein of interest into the patient, injecting a plasmid encoding for the protein of interest into the patient, injecting the protein of interest into the patient, or systemically delivering the protein of interest into the patient.
(173) In some embodiments, the patient is a human.
(174) Another aspect of the present application relates to a treatment method. This method involves contacting a cell with a vector according to the present application under conditions effective to express an RNA molecule as described herein to treat the cell.
(175) A further aspect of the present application relates to a kit comprising a vector encoding an RNA-regulated destabilization domain and a vector encoding an RNA aptamer that specifically binds to said RNA-regulated destabilization domain. Suitable RNA-regulated destabilization domains and RNA aptamers are described in detail supra.
(176) In some embodiments, the kit comprises a vector encoding tDeg and vector encoding a Pepper aptamer.
EXAMPLES
(177) The following examples are provided to illustrate embodiments of the present invention but they are by no means intended to limit its scope.
(178) Materials and Methods for Examples 1-5
(179) General methods and materials. Single stranded synthetic DNA oligonucleotides for PCR were purchased from Integrated DNA Technologies. Phusion® High-Fidelity DNA Polymerase (NEB M0530) was used for routine PCR amplifications. PCR products were run on 1% TAE agarose gels. PCR products with correct size were then excised and purified with the Qiaquick Gel Extraction kit (Qiagen 28704). Restriction endonucleases used for restriction digest were purchased from New England Biolabs, and used according to the manufacturer's recommended protocol. DNA ligation reactions were carried out using the Quick Ligation™ Kit (NEB M2200L). DNA plasmids were propagated using chemically competent *E. coli* (Agilent 200314). The QIAprep Spin Plasmid Miniprep Kit (Qiagen 27106) was used for DNA plasmid extraction and purification from *E. coli*. DNA sequencing (GENEWIZ) was used to verify the inserted gene sequences.
(180) Cell culture and transfection. HEK293T/17 (ATCC CRL-11268), U2OS (ATCC HTB-96), COS-7 (ATCC CRL-1651), and HeLa (ATCC CCL-2) cells were cultured in DMEM (Thermo Fisher Scientific 11995-065) supplemented with 10% fetal bovine serum (Corning 35-010-CV), 100 U ml.sup.−1 penicillin and 100 μg ml.sup.−1 of streptomycin (Thermo Fisher Scientific 15140122) under 37° C. with 5% CO.sub.2. TrypLE Express (Thermo Fisher Scientific 12604013) was used for detaching cells from culture flasks during cell passage. All cell lines used in this study were transfected using FuGENE HD (Promega 2311) according to the manufacturer's instructions. Prior to live-cell imaging, cells were changed to imaging media: phenol red-free DMEM (Thermo Fisher Scientific 31053-028) supplemented with 10% fetal bovine serum (Corning 35-010-CV), 100 U ml.sup.−1 penicillin and 100 μg ml.sup.−1 of streptomycin (Thermo Fisher Scientific 15140122), 1× GlutaMAX™ (Thermo Fisher Scientific 35050-061), and 1 mM sodium pyruvate (Thermo Fisher Scientific 11360-070).
(181) Fluorescence and bioluminescence imaging of tDeg-tagged proteins. To construct an expression vector for EYFP, EYFP-tDeg, mNeonGreen-tDeg, mCherry-tDeg, NanoLuc-tDeg, EGFP-TetR-tDeg, EGFP-EZH2-tDeg, or mCherry-NF-κB-tDeg, a pcDNA3.1(+) vector was digested by MluI and XbaI and ligated to an insert comprising a miniCMV promoter (5′-GGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAG AGCT-3′ (SEQ ID NO: 93), a HindIII

restriction site, a Kozak sequence (5′-GCCACC-3′), and the gene encoding EYFP, EYFP, mNeonGreen, mCherry, NanoLuc, EGFP-TetR, EGFP-EZH2, or mCherry-NF-κB, respectively, fused with tDeg. These expression vectors were called miniCMV-EYFP, miniCMV-EYFP-tDeg, miniCMV-mNeonGreen-tDeg, miniCMV-mCherry-tDeg, miniCMV-NanoLuc-tDeg, miniCMV-EGFP-TetR-tDeg, miniCMV-EGFP-EZH2-tDeg, and miniCMV-mCherry-NF-κB-tDeg respectively. For control constructs of miniCMV-EGFP-TetR, miniCMV-EGFP-EZH2, and miniCMV-mCherry-NF-κB, a stop codon was inserted on the immediate upstream of the coding sequence of tDeg using QuikChange Site-Directed Mutagenesis Kits (Agilent).

(182) To construct an expression vector for different circular RNAs, the Tornado expression plasmid (Litke et al., Highly Efficient Expression of Circular RNA Aptamers in Cells using Autocatalytic Transcripts," *Nat. Biotechnol.* 37:667-675 (2019), which is hereby incorporated by reference in is entirety) containing an F30 scaffold was digested, then ligated to inserts encoding the following sequences, respectively: wild-type TAR RNA (5′-GGCTCGTGTAGCTCATTAGCTCCGAGCC-3′ (SEQ ID NO: 65)), TAR Variant-1 (5′-GGCTCGTCTGAGCTCATTAGCTCCGAGCC-3′(SEQ ID NO: 67)), Pepper (TAR Variant-2) (5′-GGCTCGTTGAGCTCATTAGCTCCGAGCC-3′(SEQ ID NO: 69), or a control RNA, the MS2 hairpin (5′-ACATGAGGATCACCCATGT-3′(SEQ ID NO: 94)). These vectors were called: U6+27-tnd-wildtype TAR, TAR Variant-1, Pepper (TAR Variant-2), control RNA, respectively.

(183) For live-cell imagining experiments with HEK293T cells, HEK293T cells were seeded into 12-well flat bottom cell culture plates (Corning™ 3513) with 2×10.sup.5 cells per well, and were cultured overnight. On the next day, cells were transfected using FuGENE HD (Promega 2311) according to the manufacturer's instructions. Specifically, for imaging experiments in FIGS. **1**A-C, 550 ng of miniCMV-EYFP-tDeg were cotransfected with 550 ng of U6+27-tnd-wildtype TAR, TAR Variant-1, Pepper (TAR Variant-2), or control RNA, respectively. In the case of EYFP, 550 ng of miniCMV-EYFP was transfected with 550 ng of diluent DNA (pUC19 plasmid) to maintain 1.1 μg of total plasmid DNA per well. For imaging experiments in FIGS. **6**A-**6**G and FIGS. **7**A-**7**G, 550 ng of miniCMV-protein X-tDeg (protein X=mNeonGreen, mCherry, NanoLuc, EGFP-TetR, EGFP-EZH2, or mCherry-NF-κB) was cotransfected with 550 ng of circular Pepper (TAR Variant-2) or with 550 ng of diluent DNA (pUC19 plasmid). At 24 hours after transfection, cells were subcultured into 35 mm imaging dishes precoated with poly-D-lysine (Mattek Corporation P35GC-1.5-14C) and mouse laminin I (Cultrex® 3401-010-02) in culture media. Cells were then cultured overnight. Cell culture media was changed imaging media prior to fluorescence or bioluminescence live-cell imaging.

(184) For live-cell imagining experiments in FIGS. **4**A-**4**B, U2OS cells, COS-7 cells, or HeLa cells were seeded into 35 mm imaging dishes precoated with poly-D-lysine (Mattek Corporation P35GC-1.5-14C) with 2×10.sup.5 cells per dish, respectively. On the next day, cells were transfected using FuGENE HD (Promega 2311) according to the manufacturer's instructions. Specifically, 1.4 μg of miniCMV-EYFP-tDeg was cotransfected with 1.4 μg of circular Pepper (TAR Variant-2) or 1.4 μg of diluent DNA (pUC19 plasmid). At 48 hours after transfection, cell culture media was changed imaging media prior to fluorescence live-cell imaging.

(185) Prior to live-cell fluorescence or bioluminescence imaging, 1 μL of Hoechst 33342 (Thermo Fisher Scientific H3570) per 2 ml of imaging media was added to the cells. In the case of proteasome inhibitor treatment, cells were treated with either DMSO or 10 μM (final concentration in the media) MG132 for 7 hours prior to live-cell imaging. In the case of bioluminescence imaging of NanoLuc, 20 μL of furimazine (Promega Nano-Glo® Luciferase Assay System) per 2 ml of imaging media was added to the cells prior to bioluminescence imaging.

(186) For live-cell fluorescence or bioluminescence imaging, an epifluorescence inverted microscope (Nikon Eclipse TE2000-E) equipped with a CoolSnap HQ2 CCD camera and a 130-W Nikon mercury lamp was used. The NIS-Elements Advanced Research software (Nikon) was used to control the microscope and camera. Cells were imaged with a 20×/0.75-NA (numerical aperture)

or a 40×/0.75-NA air objective (Nikon) at 37° C. A FITC filter cube (with excitation filter 470±20 nm, dichroic mirror 495 nm (long pass), and emission filter 525±25 nm) was used for detecting EGFP-TetR-tDeg or EGFP-EZH2-tDeg with an exposure time of 500 msec. A YFP filter cube (with excitation filter 500±12 nm, dichroic mirror 520 nm (long pass), and emission filter 542±13.5 nm) was used for detecting EYFP, EYFP-tDeg, or mNeonGreen-tDeg with an exposure time of 500 msec. A TRITC filter cube (with excitation filter 560±20 nm, dichroic mirror 585 nm (long pass), and emission filter 630±37.5 nm) was used for detecting mCherry-tDeg, or mCherry-NF-κB-tDeg with an exposure time of 500 msec. A filter cube (with emission filter 460±25 nm) was used for detecting the bioluminescence of NanoLuc with an exposure time of 3 minutes. A DAPI filter cube (with 350±25 nm excitation filter, 400 nm (long pass) dichroic mirror, and 460±25 nm emission filter) was used for detecting the Hoechst-stained nuclei in cells with an exposure time of 100-500 msec. All filters used in these filter cubes are purchased from Chroma Technology. Cell fluorescence/bioluminescence was calculated using ImageJ by measuring the mean fluorescence/bioluminescence signal in a cell's area and subtracting background based on average signal of culture media. Normalized fluorescence/bioluminescence was calculated by dividing the cell fluorescence/bioluminescence intensity of each cell to the averaged cell fluorescence/bioluminescence of the whole cell population.

(187) RT-qPCR. Total RNA was isolated from cells using Trizol according to the manufacturer's instruction. To remove residual DNA contaminations, the purified RNA was treated with DNaseI (Thermo-Fisher) according to the manufacturer's instructions. The same amount of DNaseI-treated RNA was reverse transcribed to cDNA using SuperScript IV First-Strand kit (Invitrogen) with random hexamers according to the manufacturer's instructions. To measure relative expression levels of the RNAs of interest, qPCR measurements were performed using the iQ SYBR Green Supermix with 0.250 ng of cDNA in the final reaction mix. For the amplification, the following protocol was used: 98° C. for 2 minutes, 40 cycles of 95° C. for 10 seconds, 60° C. for 40 seconds. Primer sets for amplifying the cDNA of EYFP and mCherry are listed in Table 9. Every primer set was tested for its efficiency. To test primer specificity, melting curves were performed at the end of the 40 cycles of amplification. In the case of mCherry quantification, an untransfected sample was added as additional negative control. Relative measurements (2{circumflex over ( )}-ΔCq) of mCherry, EYFP were performed using GAPDH and RPS18 as housekeeping genes. Biological replicates were tested.

(188) TABLE-US-00015 TABLE 9 ssDNA oligo probes used in RT-qPCR EYFP fw ACGTAAACGGCCACAAGTTC SEQ ID NO: 95 EYFP rv CTTCATGTGGTCGGGGTAGC SEQ ID NO: 96 mCherry fw CACGAGTTCGAGATCGAGGG SEQ ID NO: 97 mCherry rv CAAGTAGTCGGGGATGTCGG SEQ ID NO: 98

(189) Gel staining. Total RNA was isolated from cells using TRIzol® according to the manufacturer's instruction. Then, 2.5 μg of isolated total RNA was separated using a precast 6% TBE-Urea Gel (Life Technologies EC68655). This gel was run at 200 V in TBE buffer until completion, and stained with SYBR Gold (ThermoFisher S11494) diluted 1:10,000 in TBE buffer for 15 minutes. After SYBR Gold staining, RNA bands were imaged on a ChemiDoc XRS+ system (Bio-Rad).

(190) mRNA imaging using tDeg and Pepper. To construct an expression vector for RNA-regulated fluorescent fusion proteins used in mRNA imaging, a pcDNA3.1(+) vector was digested by MluI and XbaI and ligated to an insert comprising a miniCMV promoter (5′-GGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAG CT-3′ (SEQ ID NO: 118)), a HindIII restriction site, a Kozak sequence (5′-GCCACC-3′), and the gene encoding tandem copies of mNeonGreen, mVenus, or mCherry, respectively. To construct an expression vector for an mCherry mRNA reporter containing different 3′UTR tags comprising 10 or 20 concatenated Pepper aptamers, a pcDNA3.1(+) vector was first digested by HindIII and XbaI and ligated to an insert

encoding the gene of mCherry followed by XhoI after its stop codon. This vector was called CMV-mCherry. CMV-mCherry was then digested XhoI and XbaI, and ligated to different Pepper tags, respectively. All the Pepper tags were synthesized by GenScript.
(191) U2OS cells were seeded into 35 mm imaging dishes precoated with poly-D-lysine (Mattek Corporation P35GC-1.5-14C) with 2×10.sup.5 cells per dish. On the next day, cells were transfected using FuGENE HD (Promega 2311) according to the manufacturer's instructions. Specifically, 1.4 μg of RNA-regulated fluorescent fusion protein plasmids were cotransfected with 1.4 μg of mRNA reporter plasmids. At 48 hours after transfection, cell culture media was changed to imaging media prior to imaging experiments.
(192) For mRNA imaging experiments, an epifluorescence inverted microscope (Olympus IX-70) equipped with a Evolve® 512 EMCCD OEM camera (Photometrics) and an Insight SSI 7 color solid state illumination system (Applied Precision) was used. The Resolve3D softWoRx-Acquire Version: 6.5.2 was used to control the microscope and camera. Cells were imaged with a 100×/1.4-NA oil objective at 37° C., with N=1.520 immersion oil (Applied Precision). A FITC filter cube (with excitation filter 475±14 nm, dichroic mirror with a reflection band of 481-502 nm, and a transmission band of 506-543 nm), and emission filter 525±25 nm) was used for detecting mNeonGreen with an exposure time of 50 msec. A YFP filter cube (with excitation filter 513±8.5 nm, dichroic mirror with a reflection band of 496-528 nm, and a transmission band of 537-550 nm, and emission filter 559±19 nm) was used for detecting mVenus with an exposure time of 100 msec. A TRITC filter cube (with excitation filter 542±13.5 nm, dichroic mirror with a reflection band of 547-565 nm, and a transmission band of 576-630 nm, and emission filter 594±22.5 nm) was used for detecting reporter plasmids encoding mCherry with an exposure time of 10-100 msec. Signal-to-noise ratio of the fluorescent puncta was calculated by the mean fluorescence intensity of each mRNA puncta divided by the mean fluorescence intensity of the adjacent cytosolic background fluorescence.
(193) Northern blot. HEK293T cells were seeded into 10 cm culture dish with 3×10.sup.6 cells per dish. On the next day, cells were cotransfected with CMV-mCherry-(F30-2×Pepper).sub.10 and miniCMV-(mNeonGreen).sub.4-tDeg or pUC19, respectively. A total amount of 19 μg plasmid DNA was used for each culture dish, and pUC19 vector was used here as a diluent DNA to ensure the same amount of plasmid DNA transfected to the cells. All transfections were performed using FuGENE HD (Promega 2311) according to the manufacturer's instructions. Cells were harvested after 48 hours of transfection. Total RNA was extracted with TRIzol® (Thermo Fisher Scientific 15596026) followed by isopropanol precipitation. The purified total RNA was then subjected to RNase-free DNase I (Thermo Fisher Scientific AM2224) digestion at 37° C. for 1 hour. After digestion, the RNA was subjected to phenol-chloroform (Thermo Fisher Scientific AM9720) extraction and ethanol purification.
(194) For gel electrophoresis, a 1.5% agarose/formaldehyde gel (20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA, 1.5% w/v agarose, 2% formaldehyde) was used. 20 μg of total RNA was loaded in each lane. The RNA was resuspended in 20 μL of RNA sample buffer (20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA, 50% v/v formamide, 3.7% formaldehyde). The RNA samples were heated at 70° C. for 10 minutes, and then chilled on ice for more than 1 minute. Before loading the RNA samples into the gel, the RNA samples were mixed with 2 μL of loading buffer (50% glycerol, 5 mM EDTA, 0.4% bromophenol blue, 0.4% xylene cyanol). The gel was run at 70 V for 2 hours. After electrophoresis, the gel was stained with 1×SYBR™ Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific S11494) to assess the quality of the RNA and check for separation. All solutions mentioned above were made in diethylpyrocarbonate (DEPC)-treated water.
(195) After electrophoresis, the RNA was transferred to Amersham Hybond-N+ nylon membrane (GE Healthcare Life Sciences RPN203B) using the VacuGene XL Vacuum Blotting System (GE Healthcare Life Sciences) according to the manufacturer's instructions. The RNA was then UV crosslinked to the nylon membrane. The membrane was washed with NorthernMax®

Prehybridization/Hybridization Buffer (Thermo Fisher Scientific AM8677) at 42° C. for at least 30 minutes. Biotinylated (at 5′) single-stranded DNA probes (Integrated DNA Technologies) as shown in Table 10 were mixed with NorthernMax® Prehybridization/Hybridization Buffer and incubated with the membrane at 42° C. overnight. On the following day, the membrane was washed in 50 mL of wash buffer 1 (2×SSC, 0.1% SDS) twice at 42° C. for 10 minutes each time, and then washed with wash buffer 2 (0.1×SSC, 0.1% SDS) twice at 42° C. for 15 minutes. The membrane was visualized by Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific 89880).

(196) TABLE-US-00016 TABLE 10 ssDNA oligo probes used in FIG. 12A Probe-1 GTTGAGTGATTAGCGATTGA SEQ ID NO: 99 TTCCGGCC Probe-2 GTCGGATGATTTTCGTAATA SEQ ID NO: 100 GATTGCGCTG Probe-3 TTGACGTGATTTTGTGAGAT SEQ ID NO: 101 TTTCCGCAG Probe-4 TGCCTGATTGTAAGTATGTG SEQ ID NO: 102 GATTATCGG Probe-5 GGATAGGTATGGAGGAAGTA SEQ ID NO: 103 GCTTGGA Probe-6 ACAATATCTTGCGCCGTTCG SEQ ID NO: 104 ATCTTG Probe-7 GGCCGCCAAGAAGAACGACC SEQ ID NO: 105 AA Probe-8 CCTAAGAACCTAACATATCT SEQ ID NO: 106 AGCGAGG Probe-9 TGTGCACCTTGAAGCGCATGAA SEQ ID NO: 107 Probe-10 CCTGGGTCACGGTCACCACG SEQ ID NO: 108 Probe-11 GCCCATGGTCTTCTTCTGC SEQ ID NO: 109 Probe-12 GGGTGCTTCACGTAGGCCTT SEQ ID NO: 110 Probe-13 GTCACCTTCAGCTTGGCGGTC SEQ ID NO: 111 Probe-14 GCCTCTGCTTGATCTCGCCCTTC SEQ ID NO: 112 Probe-15 GTCTTGACCTCAGCGTCGTAGTG SEQ ID NO: 113 Probe-16 CGGCGCGTTCGTACTGTTCC SEQ ID NO: 114 Probe-17 GCCGATAATCCACATACTTACAA SEQ ID NO: 115 TCAGG

(197) Imaging membrane-tethered mRNA. U2OS cells were seeded 72 hours before imaging in 96-well glass bottom dishes (Matriplates, Brooks Life Science Systems) at 40% confluency. Cells were transfected with DNA plasmids that encode miniCMV-(mNeonGreen).sub.4-tDeg, PCP-3×mCherry-CAAX and the mRNA reporter 48 hours before imaging using 0.5 μl FuGENE 6 (Promega) and 200-300 ng DNA per well. The transfection mix was prepared in OptiMEM (Sigma-Aldrich) and added to the cells in a total volume 150-200 μl of medium.

(198) Twenty-four hours prior to imaging, transcription of the reporters was induced by addition of doxycycline (1 ng/ml) (Sigma-Aldrich). Thirty minutes before imaging, the cell culture medium was replaced with pre-warmed CO.sub.2-independent Leibovitz's-15 medium (Gibco) with doxycycline. Images were acquired using a Nikon TI inverted microscope with perfect focus system equipped with a Yokagawa CSU-X1 spinning disc, a 100× 1.49 NA objective and an iXon Ultra 897 EMCCD camera (Andor) and was controlled by NIS software (Nikon). During the experiment, cells were maintained at a constant temperature of 37° C. Single Z-plane images were acquired, with the bottom plasma membrane of the cell in the focal plane. Camera exposure times of 500 ms were used for both mNeonGreen and mCherry.

(199) To determine the fluorescence intensity of mRNA foci, mean spot intensities were measured in Image J in a region of interest (ROI) 0.53×0.53 μm in size. For each spot, local background fluorescence intensity was measured in a ROI (0.53×0.53 μm in size) directly next to the spot of interest, and mean background fluorescence intensities were subtracted from the mean spot intensity. Cells with very high number of mRNAs (more than ˜50) were excluded from the analysis.

(200) Western Blotting. Cells were lysed in whole cell lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM EDTA, 50 mM NaCl, 1% Triton X-100, 0.1% SDS) containing 1× protease and phosphatase inhibitor (Pierce, 78440). Lysates were cleared by centrifugation (12,000 g for 10 minutes). Protein quantification was performed using the Pierce BCA protein assay kit according to the manufacturer's instruction (Thermo Fisher Scientific, 23227). Equal quantities of proteins were

mixed with loading dye, and incubated at 95° C. for 5 minutes before they were separated on 4-12% Bis-Tris gels (Invitrogen) and transferred onto a PVDF membrane at constant 350 mA at 4° C. for 1 hour. Membranes were blocked by incubation in 5% milk for 1 hour at room temperature under agitation and then incubated with the following primary antibodies: mouse anti-GAPDH (Santa Cruz) with a 1:5000 dilution in 1% milk overnight, or rabbit anti-mCherry (Abcam, ab167453) with a 1:1000 dilution in 1% milk overnight, or rabbit anti-ubiquitin (Abcam, ab19247) with a 1:1000 dilution in 1% milk overnight. After incubation with the appropriate secondary antibodies conjugated to HRP and extensive washing, blots were imaged on a ChemiDoc XRS+ system (Bio-Rad).

(201) Imaging ER-targeting mRNA. To construct an expression vector for an ER-targeting mRNA reporter, DNA sequence that encodes the first 29 amino acids of cytochrome p450, CytERM, and a linker sequence (MDPVVVLGLCLSCLLLLSLWKQSYGGGKLGGSGGTGGSGTSGG (SEQ ID NO: 116) was cloned into the upstream of the mCherry sequence of the CMV-mCherry-(F30-2×Pepper).sub.10 plasmid to make CMV-CytERM-mCherry-(F30-2×Pepper).sub.10. To construct the plasmid that encodes the RNA-regulated fluorescent fusion protein used in this experiment, the miniCMV promoter sequence in miniCMV-(mNeonGreen).sub.4-tDeg was replaced with the human ubiquitin C promoter sequence to make UbC-(mNeonGreen).sub.4-tDeg.

(202) U2OS cells were seeded into 35 mm imaging dishes precoated with poly-D-lysine (Mattek Corporation P35GC-1.5-14C) with 2×10.sup.5 cells per dish. On the following day, cells were cotransfected with 1.4 μg of CMV-CytERM-mCherry-(F30-2×Pepper).sub.10, 0.28 μg of UbC-(mNeonGreen).sub.4-tDeg, and 1.12 μg of pUC19 (as a diluent DNA) using FuGENE HD (Promega 2311) according to the manufacturer's instructions. At 48 hours after transfection, cell culture media was changed to imaging media prior to imaging experiments. This imaging setup for these experiments are the same as the one used for mRNA imaging using tDeg and Pepper.

(203) Imaging β-actin mRNA after arsenite stress. To construct an expression vector for a β-actin mRNA reporter containing a (F30-2×Pepper).sub.10 tag, the full length β-actin gene (from Addgene Plasmid #27123) was amplified by PCR and digested by XhoI and HindIII, and then ligated to a vector from CMV-mcherry-(F30-2×Pepper).sub.10 digested by the same restriction endonucleases to cut out the gene encoding mCherry. This expression vector was called CMV-O-actin-(F30-2×Pepper).sub.10.

(204) U2OS cells stably expresses Halo-G3BP1 were seeded into 35 mm imaging dishes precoated with poly-D-lysine (Mattek Corporation P35GC-1.5-14C) with 2×10.sup.5 cells per dish. On the following day, cells were cotransfected with 1.4 μg of miniCMV-(mNeonGreen).sub.4-tDeg with 1.4 μg of CMV-O-actin-(F30-2×Pepper).sub.10 using FuGENE HD (Promega 2311) according to the manufacturer's instructions. For control experiments, 1.4 μg of miniCMV-(mNeonGreen).sub.4-tDeg with 1.4 μg of U6+27-tnd-Pepper was used following the same transfection protocol. At ˜40 hours after transfection, cell culture media was changed to imaging media with the HaloTag® TMRDirect™ Ligand (Promega G2991) for 5 hours. Cells were then rinsed with 1×PBS (Thermo Fisher Scientific 10010049) and incubated in imaging media prior to imaging experiments. The same microscope setup as in the above mRNA imaging experiments was used. To induce stress granule formation, 1 mL of imaging media supplemented with 1000 μM of sodium arsenite was added to the cells cultured in 1 mL of imaging media to reach a final concentration of 500 μM of sodium arsenite.

(205) Statistical analysis. All data were expressed as means±s.d. with sample sizes (n) listed for each experiment. Statistical analyses were performed using Excel (Microsoft) and Prism (Graphpad). For different circular TAR variants' inhibition of tDeg's destabilizing effect, and optimization of the number of fluorescent mNeonGreen monomers in the RNA-regulated fluorescent fusion protein for imaging mRNA in live cells, one-way ANOVA was used to analyze significant differences between group means. For Pepper RNA-dependent regulation of protein stability, imaging green Pepper-tagged β-actin mRNA, proteasomal inhibition, imaging membrane-

tethered mRNA, two tailed Student's t-tests were used to analyze significant differences between group means. P values were reported for each experiment.

Example 1—tDeg Reduces Protein Stability by Inducing Proteasomal Degradation

(206) In order to expand fluorescent aptamer-based imaging, Applicant sought to create a new class of RNA-regulated fluorescent dyes that are genetically encoded. Fluorescent proteins are particularly useful since a diverse array of spectrally distinct proteins have been described (Rodriguez et al., "The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins," *Trends Biochem. Sci.* 42:111-129 (2017), which is hereby incorporated by reference in its entirety). However, these proteins are constitutively fluorescent. To make them dependent on RNA, Applicant considered making them rapidly degraded in cells except when bound by a specific RNA aptamer. In this way, fluorescence would be selectively associated with RNA-protein complexes, and not with unbound fluorescent protein. This would be functionally equivalent to RNA-induced fluorescence of small molecule dyes.

(207) First, a "destabilization domain" that can be inhibited by an RNA aptamer was developed. Previously, the Arg-Arg-Arg-Gly (SEQ ID NO: 117) was described as a degron sequence when appended to the C-terminus of proteins (Bonger et al., "Small-Molecule Displacement of a Cryptic Degron Causes Conditional Protein Degradation," *Nat. Chem. Biol.* 7:531-537 (2011), which is hereby incorporated by reference in its entirety). This sequence is similar to the arginine-rich RNA-binding domain of the Tat protein, which contains Arg-Arg as its last two amino acids. Therefore, Arg-Gly was appended to extend this Arg-Arg sequence so that the full Arg-Arg-Arg-Gly (SEQ ID NO: 117) degron is at the C-terminus of this peptide (FIGS. **1**A-**1**B and FIGS. **2**A-**2**B). This 19-amino acid-long bifunctional peptide was termed "tDeg." Tat binds a 28 nt-long RNA hairpin termed TAR (Ye et al., "Molecular Recognition in the Bovine Immunodeficiency Virus Tat Peptide-TAR RNA Complex," *Chem. Biol.* 2:827-40 (1995) and Puglisi et al., "Solution Structure of a Bovine Immunodeficiency Virus Tat-TAR Peptide-RNA Complex," *Science* 270:1200-1203 (1995), which are hereby incorporated by reference in their entirety), which may shield the degron and thus prevent recruitment of the proteasomal machinery needed for proteolysis (FIG. **1**A and FIGS. **2**A-**2**B).

(208) Whether tDeg confers instability to proteins was first investigated. To do so, tDeg was fused to the C-terminus of enhanced yellow fluorescent protein (EYFP), and the resulting fusion protein (EYFP-tDeg) was expressed in HEK293T cells. While EYFP was readily detectable, EYFP-tDeg was nearly undetectable (FIGS. **1**B-**1**C). EYFP-tDeg was restored by proteasome inhibition (FIGS. **3**A-**3**B) indicated that tDeg reduces protein stability by inducing proteasomal degradation.

Example 2—tDeg is Regulated by TAR RNA and TAR RNA Variants

(209) Whether the tDeg can be regulated by the TAR RNA was next investigated. The TAR RNA was expressed as a circular RNA using the Tornado ribozyme-assisted circularization approach to achieve high expression in mammalian cells (Litke & Jaffrey, "Highly Efficient Expression of Circular RNA Aptamers in Cells Using Autocatalytic Transcripts," *Nat. Biotechnol.* 37:667-675 (2019), which is hereby incorporated by reference in its entirety). When TAR was expressed, EYFP-tDeg-expressing cells exhibited a 24-fold increase of fluorescence relative to control RNA (FIGS. **1**B-**1**C). TAR variants that bind Tat with higher affinity, Variant-1 and Variant-2 (Smith et al., "Altering the Context of an RNA Bulge Switches the Binding Specificities of Two Viral Tat Proteins," *Biochemistry* 37:10808-10814 (1998), which is hereby incorporated by reference in its entirety), were even more efficient at inducing EYFP-tDeg, with Variant-2 exhibiting a 38-fold increase in cellular fluorescence (FIGS. **1**B-**1**C; FIGS. **4**A-**4**B). Expression of Variant-2 induced EYFP-tDeg cellular fluorescence levels similar to levels in cells expressing EYFP without the tDeg (FIG. **1**C). Furthermore, Variant-2 induced EYFP-tDeg fluorescence in diverse cell types (FIGS. **5**A-**5**G). Thus, the EYFP-tDeg is a RNA-regulated fluorescent fusion protein that is regulated by TAR.

(210) Because the TAR Variant-2 aptamer can control the expression of different colored

fluorescent proteins, as described infra, this aptamer was named after the multicolored vegetable Pepper, in keeping with the vegetable nomenclature system used previously for fluorogenic RNA aptamers.

Example 3—tDeg Tag is a Versatile Tag for Pepper-Dependent Protein Stabilization

(211) Whether the expression level of other proteins could be controlled by the Pepper RNA was next investigated. Addition of tDeg to the C-terminus of mNeonGreen, mCherry, NanoLuc, tetracycline repressor protein (TetR), EZH2, and NF-κB, resulted in minimal or undetectable protein levels in control cells and clear induction in circular Pepper-expressing cells (FIGS. **6**A-**6**G and FIGS. **7**A-**7**G). Taken together, these data indicate that the tDeg tag is a versatile tag for RNA-dependent protein stabilization.

Example 4—Intracellular Imaging Using Pepper-Modified mRNA

(212) mRNAs are commonly imaged using tethered fluorescent proteins. For example, a GFP fusion with MS2 phage coat protein (MCP) can be recruited to mRNAs containing 24-48 consecutive MS2 RNA hairpins in their 3′UTRs (Bertrand et al., "Localization of ASH1 mRNA Particles in Living Yeast," *Mol. Cell* 2:437-45 (1998), which is hereby incorporated by reference in its entirety). In this way, many GFPs are recruited to single mRNAs resulting in an aggregate fluorescence that can be detected by fluorescence microscopy. Typically nuclear localization elements are added to the GFP-MCP fusion to remove the unbound fluorescent protein from the cytoplasm into the nucleus (Bertrand et al., "Localization of ASH1 mRNA Particles in Living Yeast," *Mol. Cell* 2:437-45 (1998), which is hereby incorporated by reference in its entirety). This can reduce the fluorescence background in the cytosol, facilitating mRNA detection. However, this may introduce a potential artifact since the MS2-tagged mRNAs will contain dozens of nuclear localization sequences due to the recruited fluorescent proteins (Tyagi, S., "Imaging Intracellular RNA Distribution and Dynamics in Living Cells," *Nat. Methods* 6:331-338 (2009), which is hereby incorporated by reference in its entirety). The RNA aptamers described herein do not introduce a cellular trafficking element and may therefore bypass this concern.

(213) To investigate the use of RNA aptamers in intracellular imaging, a tag for mRNA imaging consisting of consecutive Pepper aptamers was next generated. In optimization experiments, an mCherry mRNA reporter containing different 3′UTR tags comprising 10 or 20 concatenated Pepper aptamers and Pepper aptamers that were inserted into an RNA three-way junction sequence termed F30 were imaged. Aptamers inserted within the F30 show improved folding (Filonov et al., "In-Gel Imaging of RNA Processing Using Broccoli Reveals Optimal Aptamer Expression Strategies," *Chem. Biol.* 22:649-60 (2015), which is hereby incorporated by reference in its entirety). mCherry mRNA was readily detectable as mobile fluorescent puncta in the cytoplasm when the tag contained 20 Pepper aptamers. The brightest puncta were seen when using the (F30-2×Pepper).sub.10 tag, which comprises 10 consecutive F30 sequences, with each of the two arms of F30 containing one Pepper aptamer (FIGS. **8**A—B; FIGS. **9**A-**9**D; and FIGS. **10**A-**10**C).

(214) mRNA imaging using RNA-regulated fluorescent fusion proteins of different brightness was also investigated. These proteins comprised 2, 3, or 4 tandem mNeonGreen monomers with a C-terminal tDeg. In these experiments, a RNA-regulated fluorescent fusion protein comprising four mNeonGreens provided the highest signal-to-noise ratio for imaging mRNAs (FIGS. **10**A-**10**C). Although most fluorescent puncta were detected in the cytoplasm, occasional puncta were detected in the nucleus, potentially reflecting mRNAs prior to nuclear export (FIGS. **11**A-**11**C).

(215) Cellular puncta likely reflect single mRNA molecules rather than Pepper-containing mRNA fragments since northern blotting of total cellular RNA derived from cells expressing (F30-2×Pepper).sub.10-tagged mRNA, either with or without coexpression of the (mNeonGreen).sub.4-tDeg showed mostly full-length transcripts (FIG. **12**A). Furthermore, puncta derived from mRNAs tagged with (F30-2×Pepper).sub.10 were the same size and intensity as mRNAs tagged using the PP7 fluorescent protein recruitment system, which was previously shown to reflect single mRNA molecules (Yan et al., "Dynamics of Translation of Single mRNA Molecules In Vivo," *Cell*

165:976-989 (2016), which is hereby incorporated by reference in its entirety) (FIGS. **12**B-**12**D).
(216) Adding the Pepper tag to an mRNA could adversely affect mRNA fate. However, the (F30-2×Pepper).sub.10 Pepper tag was not found to substantially alter the stability of the mCherry transcript (FIG. **13**A). Similarly, a significant difference in protein translation between the untagged and Pepper-tagged mCherry mRNA transcript was not observed (FIGS. **13**B-**13**D). Lastly, expression of RNA-regulated fluorescent fusion proteins did not significantly affect total cellular proteasome activity (FIG. **13**E).
(217) mRNAs that exhibit specific subcellular localizations were next imaged. mRNA localization to the endoplasmic reticulum (ER) was imaged using an ER-targeted reporter mRNA that encodes the first 29 amino acids of cytochrome P450, CytERM (cytoplasmic end of an endoplasmic reticulum signal-anchor membrane protein) (Costantini et al., "Assessing the Tendency of Fluorescent Proteins to Oligomerize Under Physiologic Conditions," *Traffic* 13:643-649 (2012), which is hereby incorporated by reference in its entirety). This sequence tethers the mRNA to the outer ER membrane during protein translation, and restricts the mRNA's mobility. Indeed, fluorescent puncta with low mobility were observed when this mRNA was expressed with a 3′UTR (F30-2×Pepper).sub.10 Pepper tag (FIGS. **14**A-**14**D). Treatment with puromycin, which disrupts the ribosome and dissociates the mRNA from the nascent peptide, significantly increased puncta mobility, consistent with dissociation of the reporter mRNA from the ER (FIGS. **14**A-**14**D).
(218) Next, β-actin mRNA containing a 3′UTR (F30-2×Pepper).sub.10 tag was expressed and its localization was imaged in response to arsenite treatment, which induces stress granule formation (Tourrière et al., "The RasGAP-Associated Endoribonuclease G3BP Assembles Stress Granules," *J. Cell Biol.* 160:823-831 (2003), which is hereby incorporated by reference in its entirety). Upon application of 500 μM arsenite, the individual fluorescent puncta rapidly accumulated to form stress granules as evidenced by coexpression of Halo-tagged G3BP1 to label stress granules (FIGS. **15**A-C and FIGS. **16**A-B).
Example 5—Imaging of Pepper-Regulated mVenus and Pepper-Regulated mCherry
(219) To expand the color palette of RNA-regulated fluorescent fusion proteins, two tandem copies of mVenus and two tandem copies of mCherry were fused with a C-terminal tDeg tag to convert them into RNA-regulated fluorescent fusion proteins, respectively, for imaging mRNAs. In both cases, fluorescent puncta were detected in the yellow and red fluorescence channels, respectively (FIGS. **17**A-**17**B). Together, these data show that Pepper-tagged mRNAs can be imaged in different colors using different fluorogenic proteins.
(220) Discussion of Examples 1-5
(221) The studies described infra demonstrate how constitutively fluorescent proteins can be converted to fluorescent proteins that are regulated by RNA aptamers. RNA-regulation was conferred to a protein by making its proteomic stability controlled by an RNA aptamer, Pepper. In this way, unbound RNA-regulated fluorescent fusion protein is rapidly degraded, but the RNA-regulated fluorescent fusionprotein bound to an specific RNA aptamer (e.g., Pepper) remains stable. Thus, these Pepper-regulated fluorescent fusion proteins are functionally analogous to RNA-regulated fluorogenic dyes. This system has the advantage of being able to use diverse fluorescent proteins with diverse spectral properties. Additionally, unlike the Spinach system (Paige et al., RNA Mimics of Green Fluorescent Protein," *Science* 333:642-646 (2011), which is hereby incorporated by reference in its entirety), the fluorescent system described herein is fully genetically encoded.
(222) Fluorophore maturation kinetics may also contribute to the low fluorescence of the Pepper system. Since the tDeg tag is highly efficient, it is possible that newly synthesized mNeonGreen is degraded prior to chromophore maturation. mNeonGreen that is bound to the RNA may persist for a sufficiently long time to mature to a fluorescent form while bound to RNA. This may further contribute to the low background fluorescence in cells.
(223) Unlike previous mRNA imaging systems, no nuclear localization elements are added to

fluorescent proteins to lower cytosolic background fluorescence. Instead, low background fluorescence is achieved by the highly efficient degradation of the unbound RNA-regulated fluorescent fusion protein. The simplicity of this system should simplify mRNA imaging.

(224) An important question is whether the tagged mRNA faithfully recapitulates behavior of the endogenous mRNA. The Pepper tag did not substantially affect the stability, translation, and localization of the specific mRNAs described herein. Nevertheless, imaging tags are best used when comparing two mRNAs that differ by a single sequence alteration, or the same mRNA compared in two different conditions. In this way the role of a putative functional RNA element or RNA-regulatory pathway can be inferred and then validated with the endogenous mRNA.

(225) Although the RNA-regulated destabilization domains were used to create fluorescent fusion proteins for RNA imaging, the ability to control protein expression levels through the Pepper aptamer can potentially enable novel synthetic biology applications. For these applications, Pepper can be expressed on its own, rather than part of an mRNA. By expressing tDeg-tagged proteins, diverse types of protein functions can be regulated by RNA aptamer expression levels.

Example 6—the tDeg-Pepper System can be Used to Selectively Modify RNA-Binding Proteins

(226) RNA-binding proteins (RBPs) bind to RNA molecules to orchestrate most biological functions in the cell. A major way to uncover previously unknown biological functions is to discover the RBPs involved in these processes. Current methods for discovering RBPs have low sensitivity. This is because current methods rely on recruiting a biotin ligase or a peroxidase to an RNA of interest to biotinylate any RBPs that are bound to this RNA. The major problem of these methods is the promiscuous activity of the biotin ligase or peroxidase would also nonspecifically biotinylate irrelevant proteins in the cytosol.

(227) To address this problem, new method for identifying RBPs with high sensitivity was developed. In this method, a biotin ligase and a peroxidase, whose activity is only turned on when it binds to the RNA target, was engineered. To achieve this, tDeg was fused to a biotin ligase, called TurboID, and an engineered peroxidase, called APEX2, respectively. The stability of these two proteins can be regulated by the Pepper RNA. This method drastically decreases the nonspecific biotinylation due to the promiscuous activity of this biotin ligase and peroxidase, thereby enabling the discovery of RBPs in living cells with high sensitivity.

(228) tDeg confers Pepper RNA-dependent regulation of a biotin ligase, TurboID, and a peroxidase, APEX2. FIG. **18**A-**18**B show that HEK293T cells transiently express EGFP-TurboID-tDeg (FIG. **18**A), and EGFP-APEX2-tDeg (FIG. **18**B), with and without the Pepper RNA aptamer, respectively. In each case, proteins were nearly undetectable unless coexpressed with the Pepper RNA. FIG. **18**C provides a schematic showing that a selectively activated biotin ligase (TurboID-tDeg) specifically biotinylates an RNA-binding protein (CELF1) that bind to the RNA sequence of interest (EDEN15). FIG. **18** D shows that TurboID-tDeg enables selective biotinylation of CELF1, while minimizing nonspecific biotinylation of proteins that do not bind to the RNA of interest (EDEN15). These results demonstrate that the tDeg-Pepper system can be used to selectively modify RNA-binding proteins.

Example 7—Tat-GG Confers Pepper RNA-Dependent Regulation

(229) Next, whether a variant of tDeg, Tat-GG, can be regulated by the Pepper RNA aptamer was examined. In these experiments, U2OS cells transiently expressed mNeonGreen-Tat-GG fusion protein with and without the circular Pepper RNA aptamer, respectively. Cells showed undetectable levels of green fluorescence without the circular Pepper RNA aptamer (FIG. **19**). The green fluorescence of mNeonGreen-Tat-GG was only restored when the circular Pepper RNA aptamer was coexpressed (FIG. **19**). Thus, these results confirm that the tDeg variant Tat-GG can be regulated by the Pepper RNA aptamer.

Example 8—HIV Tat-RRRG Confers HIV TAR-Dependent Regulation

(230) Next, whether HIV Tat-RRRG (RKKRRQRRRG; SEQ ID NO: 127) can be regulated by the HIV TAR sequence ACGAAGCUUGAUCCCGUUUGCCGGUCGAU CGCUUCGA (SEQ ID

NO: 128) was examined. In these experiments, cells transiently expressed YFP-HIV Tat-RRRG fusion protein with and without the circular HIV TAR RNA aptamer, respectively. Cells showed undetectable levels of yellow fluorescence without the circular HIV TAR RNA aptamer (FIG. **20**). The yellow fluorescence of YFP-HIV Tat-RRRG was restored when the circular HIV TAR RNA aptamer was coexpressed (FIG. **20**). Thus, these results confirm that HIV Tat-RRRG can be regulated by the HIV TAR RNA aptamer.

Claims

1. A nucleic acid molecule encoding an RNA-regulated fusion protein, said nucleic acid molecule comprising: a first nucleic acid sequence encoding a protein of interest and a second nucleic acid sequence encoding an RNA-regulated destabilization domain, wherein the second nucleic acid sequence is operably coupled to the first nucleic acid sequence, wherein the RNA-regulated destabilization domain is a bifunctional peptide comprising: a lentiviral transactivator of transcription (Tat) peptide and a degron peptide, wherein an RNA aptamer interacts with the RNA-regulated destabilization domain to stabilize the protein of interest, and wherein the RNA-regulated destabilization domain is tDeg as set forth in SEQ ID NO: 63.
2. The nucleic acid molecule according to claim 1, wherein the protein of interest is a fluorescent protein, a bioluminescent protein, an enzyme, or a transcription factor.
3. The nucleic acid molecule according to claim 1, wherein the lentiviral transactivator of transcription (Tat) peptide comprises an RNA binding site corresponding to or amino acid residues 4-17 of SEQ ID NO: 55.
4. The nucleic acid molecule according to claim 1 further comprising: a third nucleic acid sequence encoding a second protein of interest, wherein the third nucleic acid sequence is located between the first nucleic acid sequence and second nucleic acid sequence.
5. A vector comprising the nucleic acid molecule according to claim 1.
6. An expression system comprising an expression vector into which is inserted the nucleic acid molecule according to claim 1.
7. A host cell comprising the nucleic acid molecule of according to claim 1.
8. An RNA-regulated fusion protein encoded by the nucleic acid molecule according to claim 1.
9. A molecular complex comprising: an RNA-regulated fusion protein encoded by the nucleic acid molecule according to claim 1 comprising (i) a protein of interest and (ii) an RNA-regulated destabilization domain; and an RNA aptamer bound specifically to the RNA-regulated destabilization domain.
10. A host cell containing the molecular complex according to claim 9.