Toward a General Approach for RNA-Templated Hierarchical Assembly of Split-Proteins

*Jennifer L. Furman[1], Ahmed H. Badran[1], Oluyomi Ajulo[1], Jason R. Porter[1], Cliff I. Stains[1], David J. Segal[2] and Indraneel Ghosh[1]**

[1]Department of Chemistry, University of Arizona, 1306 East University Blvd, Tucson, AZ 85721

[2]University of California, Davis, Genome Center and Department of Pharmacology, Davis, CA 95616

* To whom correspondence should be addressed. E-mail: ghosh@email.arizona.edu

Supplementary Information

Supplementary Methods

Materials. Restriction enzymes, dNTPs, antarctic phosphatase, and T4 DNA ligase were purchased from New England Biolabs (NEB). Pfu Ultra polymerase was obtained from Stratagene. All DNA and RNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). RNasin, Steady-Glo® Luciferase Assay System, and Rabbit Reticulocyte Lysate (Promega or Alator) were obtained from Promega.

Cloning. The RNA binding PAZ domain of *Homo sapiens* argonaute-2 (Ago) was PCR amplified from pIRESneo-FLAG/HA Ago2 corrected (Addgene plasmid 10822),[1] which encodes residues 1-856 of *hs*Ago2. Only the RNA binding domain (residues 227-371) was amplified, since adjacent domains have endonuclease activity. This domain of Ago2 is assumed to adopt a stable structure based on the crystal structure of the corresponding domain of Ago1, which is 96% similar.[2] An additional Ago construct comprising residues 207-391 was also PCR amplified to investigate if a larger domain would be better suited for detection purposes. Ago inserts were ligated to existing plasmids containing CLuciferase and NLuciferase, generating pETDuet-CLuciferase-Ago (Sequence S1) and pETDuet-Ago-NLuciferase (Sequence S2). All methods concerning cloning of Pum2-NLuciferase and CLuciferase-Pum1 have been described.[3]

To generate the plasmid, pcDNA3.1(+)-CLuciferase-linker-NLuciferase, DNA encoding NLuciferase was PCR amplified from an existing plasmid and inserted into an unmodified pcDNA3.1(+) vector (Invitrogen). Next CLuciferase-linker(18 AA) was amplified from an existing plasmid and was ligated N-terminal to NLuciferase in the pcDNA3.1(+) vector, creating pcDNA3.1(+)-CLuciferase-linker(18 AA)-NLuciferase (Sequence S3). To generate extended linkers, CLuciferase-linker(18 AA) was PCR amplified from the existing construct using a reverse primer that encoded an additional 5 AAs C-terminal to the existing linker. This insert was then introduced in place of the initial CLuciferase-linker(18 AA) DNA, to generate a construct in which the linker was 23 AA. This technique was subsequently repeated using CLuciferase-linker(23 AA)-NLuciferase as a template to generate the 28

AA linker. To generate shorter linkers, complementary oligonucleotides encoding the linker flanked by AgeI and EcoRV restriction sites were annealed and directly ligated to the digested plasmid to replace the 18 AA linker. DNA encoding the zinc finger Aart (residues 17-190) was generated by PCR amplification from an existing plasmid and ligated into a vector containing NLuciferase, generating pETDuet-Aart-NLuciferase (Sequence S4).[4] DNA encoding the zinc finger E2C (residues 7-180) was generated by PCR amplification from an existing plasmid and ligated into a vector containing CLuciferase, generating pETDuet-CLuciferase-E2C (Sequence S5).[5] All sequences were confirmed with dideoxynucleotide sequencing at the University of Arizona Sequencing Facility.

In Vitro Transcription. Genes encoding CLuciferase-Ago, Ago-NLuciferase, CLuciferase-E2C, Aart-NLuciferase, and CLuciferase-linker-NLuciferase constructs were PCR amplified using *in vitro* transcription primers containing a T7 RNA Polymerase promoter and a KOZAK sequence in the forward primer and a 3' hairpin loop in the reverse primer. These primers were designed so that the complementary regions had melting temperatures greater than or equal to 70 °C. A typical PCR amplification included an initial heat denaturation of 95 °C for 5 min, followed by 35 cycles of heating to 95 °C, cooling at a rate of 6 °C/min to an annealing temperature of 50 °C. Elongation at 72 °C for 3 minutes completed the cycle. The PCR products were then used as templates for *in vitro* transcription, followed by protein generation in reticulocyte lysate, as described in the main text. Specific mRNA and target conditions used in each of the experiments (CLuciferase-Pum1/Pum2-NLuciferase, CLuciferase-Ago/Pum2-NLuciferase, CLuciferase-Pum1/Ago-NLuciferase, and CLuciferase-Ago/Ago-NLuciferase) are provided in Table S2.

Optimization of the Argonaute Detection Domain. Considering that the Ago domain was only successful as a detection domain when utilized in conjunction with a pumilio domain, we investigated multiple conditions to improve detection. For comparison to the *5′ guide*, we generated *5′ guide-2*, which is complementary to 16-nt at the 5' end of the RNA target (Table S1). This guide was annealed to the RNA target as described in the main text. *In vitro* translations were carried out as described, using 2

pmol each of CLuciferase-Ago and Pum2-NLuciferase or Ago-NLuciferase. The annealed targets were added at 10 nM after translation, followed by luminescence readings as described. The *5′ guide* and *5′ guide-2* provided very similar luminescence readings for both sets of sensors (Figure S1, A and B).

We additionally investigated the use of a larger Ago domain for detection purposes. Specifically, we compared Ago domains consisting of residues 227-371 and residues 207-391. *In vitro* translations were carried out as described, using 0.1 pmol each of CLuciferase-Ago and Pum2-NLuciferase or 2 pmol each of CLuciferase-Ago and Ago-NLuciferase. In the case of CLuciferase-Ago and Pum2-NLuciferase, the annealed target (RNA + *5′ guide*) was added at 10 nM after translation. In the case of CLuciferase-Ago and Ago-NLuciferase, the annealed target (RNA + *5′ guide* + *3′ guide*) was added at 10 nM during the translation. Luminescence readings were acquired as described. The two different Ago constructs provided very similar luminescence readings for both sets of sensors (Figure S1, C and D).

Generation of VEGF, HER2, and hDM2 Target mRNA. VEGF dsDNA was PCR amplified from an existing plasmid, pQE30-VEGF, which contains nucleotides 109-403 of VEGF cDNA, isoform 165. This 295 nucleotide region was amplified using the primers containing a T7 RNA polymerase promoter indicated in Table S1. HER2 dsDNA (nucleotides 785-985, based on Genbank NM_004448) was PCR amplified from an existing plasmid, pSGHV0-HER2,[6] which contains a portion of the HER2 extracellular domain, using primers indicated in Table S1. Due to the presence of contaminating PCR products, a gel extraction was performed using a QIAquick PCR purification kit (QIAGEN), resulting in isolation of a pure product, as visualized by agarose gel electrophoresis. hDM2 dsDNA nucleotides 306-854 (based on Genbank NM_002392) was amplified from an existing plasmid. *In vitro* transcription for each of these PCR products was carried out as described in the main text.

Sequence S1. DNA encoding CLuciferase-Ago. CLuciferase is in red, the 18 amino acid linker is shown in black, and Ago (residues 227-371) is in green.

```
atgtccggttatgtaaacaatccggaagcgaccaacgccttgattgacaaggatggatggctacattctggagac
 M  S  G  Y  V  N  N  P  E  A  T  N  A  L  I  D  K  D  G  W  L  H  S  G  D
atagcttactgggacgaagacgaacacttcttcatagttgaccgcttgaagtctttaattaaatacaaaggatat
 I  A  Y  W  D  E  D  E  H  F  F  I  V  D  R  L  K  S  L  I  K  Y  K  G  Y
caggtggcccccgctgaattggaatcgatattgttacaacaccccaacatcttcgacgcgggcgtggcaggtctt
 Q  V  A  P  A  E  L  E  S  I  L  L  Q  H  P  N  I  F  D  A  G  V  A  G  L
cccgacgatgacgccggtgaacttcccgccgccgttgttgttttggagcacggaaagacgatgacggaaaaagag
 P  D  D  D  A  G  E  L  P  A  A  V  V  V  L  E  H  G  K  T  M  T  E  K  E
atcgtggattacgtcgccagtcaagtaacaaccgcgaaaaagttgcgcggaggagttgtgtttgtggacgaagta
 I  V  D  Y  V  A  S  Q  V  T  T  A  K  K  L  R  G  G  V  V  F  V  D  E  V
ccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaagggcggaaag
 P  K  G  L  T  G  K  L  D  A  R  K  I  R  E  I  L  I  K  A  K  K  G  G  K
tccaaattggcctgcagggcggttcaggcggtgggggttctggcgggggtgggagccccggggcacagccagta
 S  K  L  G  L  Q  G  G  S  G  G  G  G  S  G  G  G  G  S  P  G  A  Q  P  V
atcgagtttgtttgtgaagttttggattttaaaagtattgaagaacaacaaaaacctctgacagattcccaaagg
 I  E  F  V  C  E  V  L  D  F  K  S  I  E  E  Q  Q  K  P  L  T  D  S  Q  R
gtaaagtttaccaaagaaattaaaggtctaaaggtggagataacgcactgtgggcagatgaagaggaagtaccgt
 V  K  F  T  K  E  I  K  G  L  K  V  E  I  T  H  C  G  Q  M  K  R  K  Y  R
gtctgcaatgtgacccggcggcccgccagtcaccaaacattcccgctgcagcaggagagcgggcagacggtggag
 V  C  N  V  T  R  R  P  A  S  H  Q  T  F  P  L  Q  Q  E  S  G  Q  T  V  E
tgcacggtggcccagtatttcaaggacaggcacaagttggttctgcgctacccccacctcccatgtttacaagtc
 C  T  V  A  Q  Y  F  K  D  R  H  K  L  V  L  R  Y  P  H  L  P  C  L  Q  V
ggacaggagcagaaacacacctaccttccccctggaggtctgtaacattgtggcaggacaaagatgtattaaaaaa
 G  Q  E  Q  K  H  T  Y  L  P  L  E  V  C  N  I  V  A  G  Q  R  C  I  K  K
ttaacggacaatcagacctcaaccatgatcagagcgactgctaggtcg
 L  T  D  N  Q  T  S  T  M  I  R  A  T  A  R  S
```

Sequence S2. DNA encoding Ago-NLuciferase. Ago (residues 227-371) is in green, the 17 amino acid linker is shown in black, and NLuciferase is in red.

```
gcacagccagtaatcgagtttgtttgtgaagttttggattttaaaagtattgaagaacaacaaaaacctctgaca
 A  Q  P  V  I  E  F  V  C  E  V  L  D  F  K  S  I  E  E  Q  Q  K  P  L  T
gattcccaaagggtaaagtttaccaaagaaattaaaggtctaaaggtggagataacgcactgtgggcagatgaag
 D  S  Q  R  V  K  F  T  K  E  I  K  G  L  K  V  E  I  T  H  C  G  Q  M  K
aggaagtaccgtgtctgcaatgtgacccggcggcccgccagtcaccaaacattcccgctgcagcaggagagcggg
 R  K  Y  R  V  C  N  V  T  R  R  P  A  S  H  Q  T  F  P  L  Q  Q  E  S  G
cagacggtggagtgcacggtggcccagtatttcaaggacaggcacaagttggttctgcgctacccccacctccca
 Q  T  V  E  C  T  V  A  Q  Y  F  K  D  R  H  K  L  V  L  R  Y  P  H  L  P
tgtttacaagtcggacaggagcagaaacacacctaccttcccctggaggtctgtaacattgtggcaggacaaaga
 C  L  Q  V  G  Q  E  Q  K  H  T  Y  L  P  L  E  V  C  N  I  V  A  G  Q  R
tgtattaaaaaattaacggacaatcagacctcaaccatgatcagagcgactgctaggtcgaccggtgggggtggc
 C  I  K  K  L  T  D  N  Q  T  S  T  M  I  R  A  T  A  R  S  T  G  G  G  G
ggttcaggcggtgggggttctggtgggggtggtaccgaagacgccaaaaacataaagaaaggcccggcgccattc
 G  S  G  G  G  G  S  G  G  G  G  T  E  D  A  K  N  I  K  K  G  P  A  P  F
tatcctctagaggatggaaccgctggagagcaactgcataaggctatgaagagatacgccctggttcctggaaca
 Y  P  L  E  D  G  T  A  G  E  Q  L  H  K  A  M  K  R  Y  A  L  V  P  G  T
attgcttttacagatgcacatatcgaggtgaacatcacgtacgcggaatacttcgaaatgtccgttcggttggca
 I  A  F  T  D  A  H  I  E  V  N  I  T  Y  A  E  Y  F  E  M  S  V  R  L  A
gaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactctcttcaattctttt
 E  A  M  K  R  Y  G  L  N  T  N  H  R  I  V  V  C  S  E  N  S  L  Q  F  F
atgccggtgttgggcgcgttatttatcggagttgcagttgcgcccgcgaacgacatttataatgaacgtgaattg
 M  P  V  L  G  A  L  F  I  G  V  A  V  A  P  A  N  D  I  Y  N  E  R  E  L
ctcaacagtatgaacatttcgcagcctaccgtagtgtttgtttccaaaaaggggttgcaaaaaattttgaacgtg
 L  N  S  M  N  I  S  Q  P  T  V  V  F  V  S  K  K  G  L  Q  K  I  L  N  V
caaaaaaaattaccaataatccagaaaattattatcatggattctaaaacggattaccagggatttcagtcgatg
 Q  K  K  L  P  I  I  Q  K  I  I  I  M  D  S  K  T  D  Y  Q  G  F  Q  S  M
tacacgttcgtcacatctcatctacctcccggttttaatgaatacgattttgtaccagagtcctttgatcgtgac
 Y  T  F  V  T  S  H  L  P  P  G  F  N  E  Y  D  F  V  P  E  S  F  D  R  D
aaaacaattgcactgataatgaattcctctggatctactgggttacctaagggtgtggcccttccgcatagaact
 K  T  I  A  L  I  M  N  S  S  G  S  T  G  L  P  K  G  V  A  L  P  H  R  T
gcctgcgtcagattctcgcatgccagagatcctatttttggcaatcaaatcattccggatactgcgattttaagt
 A  C  V  R  F  S  H  A  R  D  P  I  F  G  N  Q  I  I  P  D  T  A  I  L  S
gttgttccattccatcacggttttggaatgtttactacactcggatatttgatatgtggatttcgagtcgtctta
 V  V  P  F  H  H  G  F  G  M  F  T  T  L  G  Y  L  I  C  G  F  R  V  V  L
atgtatagatttgaagaagagctgtttttacgatcccttcaggattacaaaattcaaagtgcgttgctagtacca
 M  Y  R  F  E  E  E  L  F  L  R  S  L  Q  D  Y  K  I  Q  S  A  L  L  V  P
accctatttttcattcttcgccaaaagcactctgattgacaaatacgatttatctaatttacacgaaattgcttct
 T  L  F  S  F  F  A  K  S  T  L  I  D  K  Y  D  L  S  N  L  H  E  I  A  S
gggggcgcacctctttcgaaagaagtcggggaagcggttgcaaaacgcttccatcttccagggatacgacaagga
 G  G  A  P  L  S  K  E  V  G  E  A  V  A  K  R  F  H  L  P  G  I  R  Q  G
tatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtcggtaaa
 Y  G  L  T  E  T  T  S  A  I  L  I  T  P  E  G  D  D  K  P  G  A  V  G  K
gttgttccatttttttgaagcgaaggttgtggatctggataccgggaaaacgctgggcgttaatcagagaggcgaa
 V  V  P  F  F  E  A  K  V  V  D  L  D  T  G  K  T  L  G  V  N  Q  R  G  E
ttatgtgtcagaggacctatgattatgtccggttatgtaaacaatccggaagcgaccaacgccttgattgacaag
 L  C  V  R  G  P  M  I  M  S  G  Y  V  N  N  P  E  A  T  N  A  L  I  D  K
gatgga
 D  G
```

Sequence S3. CLuciferase-linker(18 AA)-NLuciferase. CLuciferase is shown in red, the 18 amino acid linker is shown in black, and NLuciferase is green. The restriction sites flanking the linker, which were used to facilitate alterations in the linker length, are boxed.

```
atgggtacctccggttatgtaaacaatccggaagcgaccaacgccttgattgacaaggatggatggctacattct
 M  G  T  S  G  Y  V  N  N  P  E  A  T  N  A  L  I  D  K  D  G  W  L  H  S
ggagacatagcttactgggacgaagacgaacacttcttcatcgttgaccgcctgaagtctctgattaagtacaaa
 G  D  I  A  Y  W  D  E  D  E  H  F  F  I  V  D  R  L  K  S  L  I  K  Y  K
ggctatcaggtggctcccgctgaattggaatccatcttgctccaacaccccaacatcttcgacgcaggtgtcgca
 G  Y  Q  V  A  P  A  E  L  E  S  I  L  L  Q  H  P  N  I  F  D  A  G  V  A
ggtcttcccgacgatgacgccggtgaacttcccgccgccgttgttgttttggagcacggaaagacgatgacggaa
 G  L  P  D  D  D  A  G  E  L  P  A  A  V  V  V  L  E  H  G  K  T  M  T  E
aaagagatcgtggattacgtcgccagtcaagtaacaaccgcgaaaaagttgcgcggaggagttgtgtttgtggac
 K  E  I  V  D  Y  V  A  S  Q  V  T  T  A  K  K  L  R  G  G  V  V  F  V  D
gaagtaccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaagggc
 E  V  P  K  G  L  T  G  K  L  D  A  R  K  I  R  E  I  L  I  K  A  K  K  G
ggaaagatcgccgtgaccggtggcggtgggggttctggcgggggtgggagcggtggcggttctgatatcgaagac
 G  K  I  A  V  T  G  G  G  G  G  S  G  G  G  G  S  G  G  G  S  D  I  E  D
gccaaaaacataaagaaaggcccggcgccattctatccgctggaagatggaaccgctggagagcaactgcataag
 A  K  N  I  K  K  G  P  A  P  F  Y  P  L  E  D  G  T  A  G  E  Q  L  H  K
gctatgaagagatacgccctggttcctggaacaattgcttttacagatgcacatatcgaggtggacatcacttac
 A  M  K  R  Y  A  L  V  P  G  T  I  A  F  T  D  A  H  I  E  V  D  I  T  Y
gctgagtacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacagaatc
 A  E  Y  F  E  M  S  V  R  L  A  E  A  M  K  R  Y  G  L  N  T  N  H  R  I
gtcgtatgcagtgaaaactctcttcaattctttatgccggtgttgggcgcgttatttatcggagttgcagttgcg
 V  V  C  S  E  N  S  L  Q  F  F  M  P  V  L  G  A  L  F  I  G  V  A  V  A
cccgcgaacgacatttataatgaacgtgaattgctcaacagtatgggcatttcgcagcctaccgtggtgttcgtt
 P  A  N  D  I  Y  N  E  R  E  L  L  N  S  M  G  I  S  Q  P  T  V  V  F  V
tccaaaaaggggttgcaaaaaattttgaacgtgcaaaaaaagctcccaatcatccaaaaaattattatcatggat
 S  K  K  G  L  Q  K  I  L  N  V  Q  K  K  L  P  I  I  Q  K  I  I  I  M  D
tctaaaacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccggttttaatgaa
 S  K  T  D  Y  Q  G  F  Q  S  M  Y  T  F  V  T  S  H  L  P  P  G  F  N  E
tacgattttgtgccagagtccttcgatagggacaagacaattgcactgatcatgaactcctctggatctactggt
 Y  D  F  V  P  E  S  F  D  R  D  K  T  I  A  L  I  M  N  S  S  G  S  T  G
ctgcctaaaggtgtcgctctgcctcatagaactgcctgcgtgagattctcgcatgccagagatcctatttttggc
 L  P  K  G  V  A  L  P  H  R  T  A  C  V  R  F  S  H  A  R  D  P  I  F  G
aatcaaatcattccggatactgcgattttaagtgttgttccattccatcacggttttggaatgtttactacactc
 N  Q  I  I  P  D  T  A  I  L  S  V  V  P  F  H  H  G  F  G  M  F  T  T  L
ggatatttgatatgtggatttcgagtcgtcttaatgtatagatttgaagaagagctgtttctgaggagccttcag
 G  Y  L  I  C  G  F  R  V  V  L  M  Y  R  F  E  E  E  L  F  L  R  S  L  Q
gattacaagattcaaagtgcgctgctggtgccaaccctattctccttcttcgccaaaagcactctgattgacaaa
 D  Y  K  I  Q  S  A  L  L  V  P  T  L  F  S  F  F  A  K  S  T  L  I  D  K
tacgatttatctaatttacacgaaattgcttctggtggcgctcccctctctaaggaagtcggggaagcggttgcc
 Y  D  L  S  N  L  H  E  I  A  S  G  G  A  P  L  S  K  E  V  G  E  A  V  A
aagaggttccatctgccaggtatcaggcaaggatatgggctcactgagactacatcagctattctgattacaccc
 K  R  F  H  L  P  G  I  R  Q  G  Y  G  L  T  E  T  T  S  A  I  L  I  T  P
gaggggggatgataaaccgggcgcggtcggtaaagttgttccatttttttgaagcgaaggttgtggatctggatacc
 E  G  D  D  K  P  G  A  V  G  K  V  V  P  F  F  E  A  K  V  V  D  L  D  T
gggaaaacgctgggcgttaatcaaagaggcgaactgtgtgtgagaggtcctatgattatgtccggttatgtaaac
 G  K  T  L  G  V  N  Q  R  G  E  L  C  V  R  G  P  M  I  M  S  G  Y  V  N
aatccggaagcgaccaacgccttgattgacaaggatgga
 N  P  E  A  T  N  A  L  I  D  K  D  G
```

Sequence S4. DNA encoding Aart-NLuciferase. Aart is in green, the 17 amino acid linker is in black, and NLuciferase is shown in red.

```
cccggggagaagccctatgcttgtccggaatgtggtaagtccttcagccgcagcgatcacctggccgaacaccag
 P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S  R  S  D  H  L  A  E  H  Q
cgtacccacacgggtgaaaaaccgtataaatgcccagagtgcggcaaatcttttagcgataagaaagatctgacc
 R  T  H  T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  D  K  K  D  L  T
cggcatcaacgcactcatactggcgagaagccatacaaatgtccagaatgtggcaagtctttcagccagcgcgca
 R  H  Q  R  T  H  T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  Q  R  A
aacctgcgcgcccaccaacgtactcacaccggggagaagccttatgcttgtccggaatgtggtaagtccttctct
 N  L  R  A  H  Q  R  T  H  T  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
cagctggcccacctgcgcgcccaccagcgtacccacacgggtgaaaaaccgtataaatgcccagagtgcggcaaa
 Q  L  A  H  L  R  A  H  Q  R  T  H  T  G  E  K  P  Y  K  C  P  E  C  G  K
tcttttagccgcgaggataacctgcacacccatcaacgtactcatactggcgagaagccatacaaatgtccagaa
 S  F  S  R  E  D  N  L  H  T  H  Q  R  T  H  T  G  E  K  P  Y  K  C  P  E
tgtggcaagtctttctcccgccgcgatgctctgaacgtgcaccaacgtactcacaccggcaaaaaaactagcacc
 C  G  K  S  F  S  R  R  D  A  L  N  V  H  Q  R  T  H  T  G  K  K  T  S  T
ggtgggggtggcggttcaggcggtgggggttctggtgggggtggtaccgaagacgccaaaaacataaagaaaggc
 G  G  G  G  G  S  G  G  G  G  S  G  G  G  G  T  E  D  A  K  N  I  K  K  G
ccggcgccattctatcctctagaggatggaaccgctggagagcaactgcataaggctatgaagagatacgccctg
 P  A  P  F  Y  P  L  E  D  G  T  A  G  E  Q  L  H  K  A  M  K  R  Y  A  L
gttcctggaacaattgcttttacagatgcacatatcgaggtgaacatcacgtacgcggaatacttcgaaatgtcc
 V  P  G  T  I  A  F  T  D  A  H  I  E  V  N  I  T  Y  A  E  Y  F  E  M  S
gttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactct
 V  R  L  A  E  A  M  K  R  Y  G  L  N  T  N  H  R  I  V  V  C  S  E  N  S
cttcaattctttatgccggtgttgggcgcgttatttatcggagttgcagttgcgcccgcgaacgacatttataat
 L  Q  F  F  M  P  V  L  G  A  L  F  I  G  V  A  V  A  P  A  N  D  I  Y  N
gaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtgtttgtttccaaaaaggggttgcaaaaa
 E  R  E  L  L  N  S  M  N  I  S  Q  P  T  V  V  F  V  S  K  K  G  L  Q  K
attttgaacgtgcaaaaaaaattaccaataatccagaaaattattatcatggattctaaaacggattaccaggga
 I  L  N  V  Q  K  K  L  P  I  I  Q  K  I  I  I  M  D  S  K  T  D  Y  Q  G
tttcagtcgatgtacacgttcgtcacatctcatctacctcccggttttaatgaatacgattttgtaccagagtcc
 F  Q  S  M  Y  T  F  V  T  S  H  L  P  P  G  F  N  E  Y  D  F  V  P  E  S
tttgatcgtgacaaaacaattgcactgataatgaattcctctggatctactgggttacctaagggtgtggcccttc
 F  D  R  D  K  T  I  A  L  I  M  N  S  S  G  S  T  G  L  P  K  G  V  A  L
ccgcatagaactgcctgcgtcagattctcgcatgccagagatcctatttttggcaatcaaatcattccggatact
 P  H  R  T  A  C  V  R  F  S  H  A  R  D  P  I  F  G  N  Q  I  I  P  D  T
gcgattttaagtgttgttccattccatcacggttttggaatgtttactacactcggatatttgatatgtggattt
 A  I  L  S  V  V  P  F  H  H  G  F  G  M  F  T  T  L  G  Y  L  I  C  G  F
cgagtcgtcttaatgtatagatttgaagaagagctgtttttacgatcccttcaggattacaaaattcaaagtgcg
 R  V  V  L  M  Y  R  F  E  E  E  L  F  L  R  S  L  Q  D  Y  K  I  Q  S  A
ttgctagtaccaaccctattttcattcttcgccaaaagcactctgattgacaaatacgatttatctaatttacac
 L  L  V  P  T  L  F  S  F  F  A  K  S  T  L  I  D  K  Y  D  L  S  N  L  H
gaaattgcttctgggggcgcacctctttcgaaagaagtcggggaagcggttgcaaaacgcttccatcttccaggg
 E  I  A  S  G  G  A  P  L  S  K  E  V  G  E  A  V  A  K  R  F  H  L  P  G
atacgacaaggatatgggctcactgagactacatcagctattctgattacacccgaggggatgataaaccgggc
 I  R  Q  G  Y  G  L  T  E  T  T  S  A  I  L  I  T  P  E  G  D  D  K  P  G
gcggtcggtaaagttgttccatttttgaagcgaaggttgtggatctggataccgggaaaacgctgggcgttaat
 A  V  G  K  V  V  P  F  F  E  A  K  V  V  D  L  D  T  G  K  T  L  G  V  N
cagagaggcgaattatgtgtcagaggacctatgattatgtccggttatgtaaacaatccggaagcgaccaacgcc
 Q  R  G  E  L  C  V  R  G  P  M  I  M  S  G  Y  V  N  N  P  E  A  T  N  A
ttgattgacaaggatgga
 L  I  D  K  D  G
```

Sequence S5. DNA encoding CLuciferase-E2C. CLuciferase is in red, the 14 amino acid linker is in black, and E2C is shown in green.

```
atgtccggttatgtaaacaatccggaagcgaccaacgccttgattgacaaggatggatggctacattctggagac
 M  S  G  Y  V  N  N  P  E  A  T  N  A  L  I  D  K  D  G  W  L  H  S  G  D
atagcttactgggacgaagacgaacacttcttcatagttgaccgcttgaagtctttaattaaatacaaaggatat
 I  A  Y  W  D  E  D  E  H  F  F  I  V  D  R  L  K  S  L  I  K  Y  K  G  Y
caggtggcccccgctgaattggaatcgatattgttacaacaccccaacatcttcgacgcgggcgtggcaggtctt
 Q  V  A  P  A  E  L  E  S  I  L  L  Q  H  P  N  I  F  D  A  G  V  A  G  L
cccgacgatgacgccggtgaacttcccgccgccgttgttgttttggagcacggaaagacgatgacggaaaaagag
 P  D  D  D  A  G  E  L  P  A  A  V  V  V  L  E  H  G  K  T  M  T  E  K  E
atcgtggattacgtcgccagtcaagtaacaaccgcgaaaaagttgcgcggaggagttgtgtttgtggacgaagta
 I  V  D  Y  V  A  S  Q  V  T  T  A  K  K  L  R  G  G  V  V  F  V  D  E  V
ccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaagggcggaaag
 P  K  G  L  T  G  K  L  D  A  R  K  I  R  E  I  L  I  K  A  K  K  G  G  K
tccaaattgggcctgcagggcggttcaggcggtgggggttctggcgggggtgggagccccggggagaagccctat
 S  K  L  G  L  Q  G  G  S  G  G  G  G  S  G  G  G  G  S  P  G  E  K  P  Y
gcttgtccggaatgtggtaagtccttcagtaggaaggattcgcttgtgaggcaccagcgtacccacacgggtgaa
 A  C  P  E  C  G  K  S  F  S  R  K  D  S  L  V  R  H  Q  R  T  H  T  G  E
aaaccgtataaatgcccagagtgcggcaaatcttttagtcagtcgggggatcttaggcgtcatcaacgcactcat
 K  P  Y  K  C  P  E  C  G  K  S  F  S  Q  S  G  D  L  R  R  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtctttcagtgattgtcgtgatcttgcgaggcaccaa
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  D  C  R  D  L  A  R  H  Q
cgtactcacaccggggagaagccctatgcttgtccggaatgtggtaagtccttctctcagagctctcacctggtg
 R  T  H  T  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S  Q  S  S  H  L  V
cgccaccagcgtacccacacgggtgaaaaaccgtataaatgcccagagtgcggcaaatcttttagtgactgccgc
 R  H  Q  R  T  H  T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  D  C  R
gaccttgctcgccatcaacgcactcatactggcgagaagccatacaaatgtccagaatgtggcaagtctttcagc
 D  L  A  R  H  Q  R  T  H  T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S
cgctctgacaagctggtgcgtcaccaacgtactcacaccggtaaaaaaactagt
 R  S  D  K  L  V  R  H  Q  R  T  H  T  G  K  K  T  S
```

Supplementary Tables

RNA guides	
5' guide	5'CUAUAUACACCAUGUU
3' guide	5'GCCGCUAUAUCAAUU
5' guide-2	5'GACUAUAUACACCAUGUU

Aart hairpin	5'GCATGTAGGGAAAAGCCCGGCGTCCTCGCCGGGCTTTTCCCTACATGC

RNA target primers	
VEGF FWD	5'GCAGCTAATACGACTCACTATAGGCATCACGAAGTGGTGAAGTTCATGGATGTCTATCAGC
VEGF REV	5'CTTTCTTTGGTCTGCATTCACATTTGTTGTGCTGTAGGAAGC
HER2 FWD	5'GCAGCTAATACGACTCACTATAGGCTGATAGACACCAACCGCTCTCGGGC
HER2 REV	5'GTGCTTGGGGCCCGTGCAGC
hDM2 FWD	5'GCAGCTAATACGACTCACTATAGGATGTGCAATACCAACATGTCTGTACCTACTGATGGTG
hDM2 REV	5'GCGTTTTCTTTGTCGTTCACCAGATAATTCATCTGAATTTTCTTCTGTCTCAC

Hairpin-guides	
VEGF(64-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAG
VEGF(81-95)-Aart	5'TCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(62-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAGGG
VEGF(81-97)-Aart	5'TCTCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(60-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAGGGTC
VEGF(81-99)-Aart	5'GATCTCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(58-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAGGGTCTC
VEGF(81-101)-Aart	5'TCGATCTCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(56-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAGGGTCTCGA
VEGF(81-103)-Aart	5'ACTCGATCTCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(54-78)-E2C	5'GAGGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGAAGATGTCCACCAGGGTCTCGATT
VEGF(81-105)-Aart	5'GTACTCGATCTCATCAGGGTACTCCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
VEGF(216-243)-E2C	5'GAGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAAGCCTTGGTGAGGTTTGATC
VEGF(237-255)-Aart	5'GCTCATCTCTCCTATGTGCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
HER2(100-118)-E2C	5'GAGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAACGGCACAGACAGTGCGCGT
HER2(122-140)-Aart	5'CCCTTGCAGCGGGCACAGCAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG
hDM2(43-61)-E2C	5'GAGGGCCGGAGCCGCAGTGCGTCCTCGCACTGCGGCTCCGGCCCCTCAAAACTGGAATCTGTGAGGTGGT
hDM2(64-82)-Aart	5'CCAGGGTCTCTTGTTCCGAAAAACTCCGGGCTTTTCCCTACATGCTCCTGCATGTAGGGAAAAGCCCGGAG

Table S1. Oligonucleotides used in this study. For the Aart hairpin, the complementary regions forming the Aart binding site are shown in red, and the loop is in green. For the RNA target primers, the T7 RNA Polymerase promoter is shown in blue. For the hairpin-guides, the complementary hairpin sequences are shown in red, the loop is in green, and the guide sequence is in blue. The numbers in parentheses

indicate the intended target binding location relative to the beginning of the transcript, which is designated as position 1. Unless otherwise specified, all VEGF experiments use the VEGF(60-78)-E2C and VEGF(81-99)-Aart guides.

Table S2. *In vitro* translation conditions used corresponding to the data presented in Figures 2A-D of the main text.

mRNA 1	mRNA 2	Final target concentration
2 pmol Pum2-NLuciferase	2 pmol CLuciferase-Pum1	10 nM target
0.1 pmol Pum2-NLuciferase	0.1 pmol CLuciferase-Ago	10 nM 5′ guide + 10 nM target
1 pmol Ago-NLuciferase	1 pmol CLuciferase-Pum1	10 nM 3′ guide + 10 nM target
2 pmol Ago-NLuciferase	2 pmol CLuciferase-Ago	10 nM 5′ guide + 10 nM 3′ guide + 10 nM target

Table S3. Data for the E2C-1-Aart detection limit corresponding to Figure 3C of the main text.

	sample 1		sample 2	
[E2C-1-Aart] (pM)	reading 1	reading 2	reading 1	reading 2
100	153	159	155.5	152.2
50	88	86	87.1	93.7
25	41	44	38.7	41.3
10	18	19	18.7	18.7
5	11	12	9.9	10.7
2	6.1	7.1	5.9	5.2
0	2.5	2.8	2.1	1.6

Table S4. Presence of guide sites in yeast and human genomes. Sites shown in blue were used in a custom algorithm (bsite, Segal and Korf, unpublished) to search for sites with up to two mismatches in the genomes of *Saccharomyces cerevisiae* (Build 2.1) and human (Build 37.1). The number of sites is shown with numbers of mismatches in parentheses.

Hairpin-guides		Sites in *Saccharomyces cerevisiae*	Sites in *Homo sapiens*
VEGF(60-78)-E2C	5'GAAGATGTCCACCAGGGTC	0	1 (0) 2 (1) 37 (2)
VEGF(81-99)-Aart	5'GATCTCATCAGGGTACTCC	0	1 (0) 4 (1) 70 (2)
VEGF(216-243)-E2C	5'GCCTTGGTGAGGTTTGATC	0	1 (0) 2 (1) 23 (2)
VEGF(237-255)-Aart	5'GCTCATCTCTCCTATGTGC	0	2 (0) 1 (1) 36 (2)
HER2(100-118)-E2C	5'CGGCACAGACAGTGCGCGT	0	1 (0) 0 (1) 2 (2)
HER2(122-140)-Aart	5'CCCTTGCAGCGGGCACAGC	0	3 (0) 0 (1) 41 (2)
hDM2(43-61)-E2C	5'CTGGAATCTGTGAGGTGGT	0	1 (0) 4 (1) 145 (2)
hDM2(64-82)-Aart	5'CCAGGGTCTCTTGTTCCGA	0	1 (0) 0 (1) 47 (2)
GAL4 (*Sc* control seq)	5'TGTTTAAACTTAAGGGATC	1 (0)	N/D

Supplementary Figures

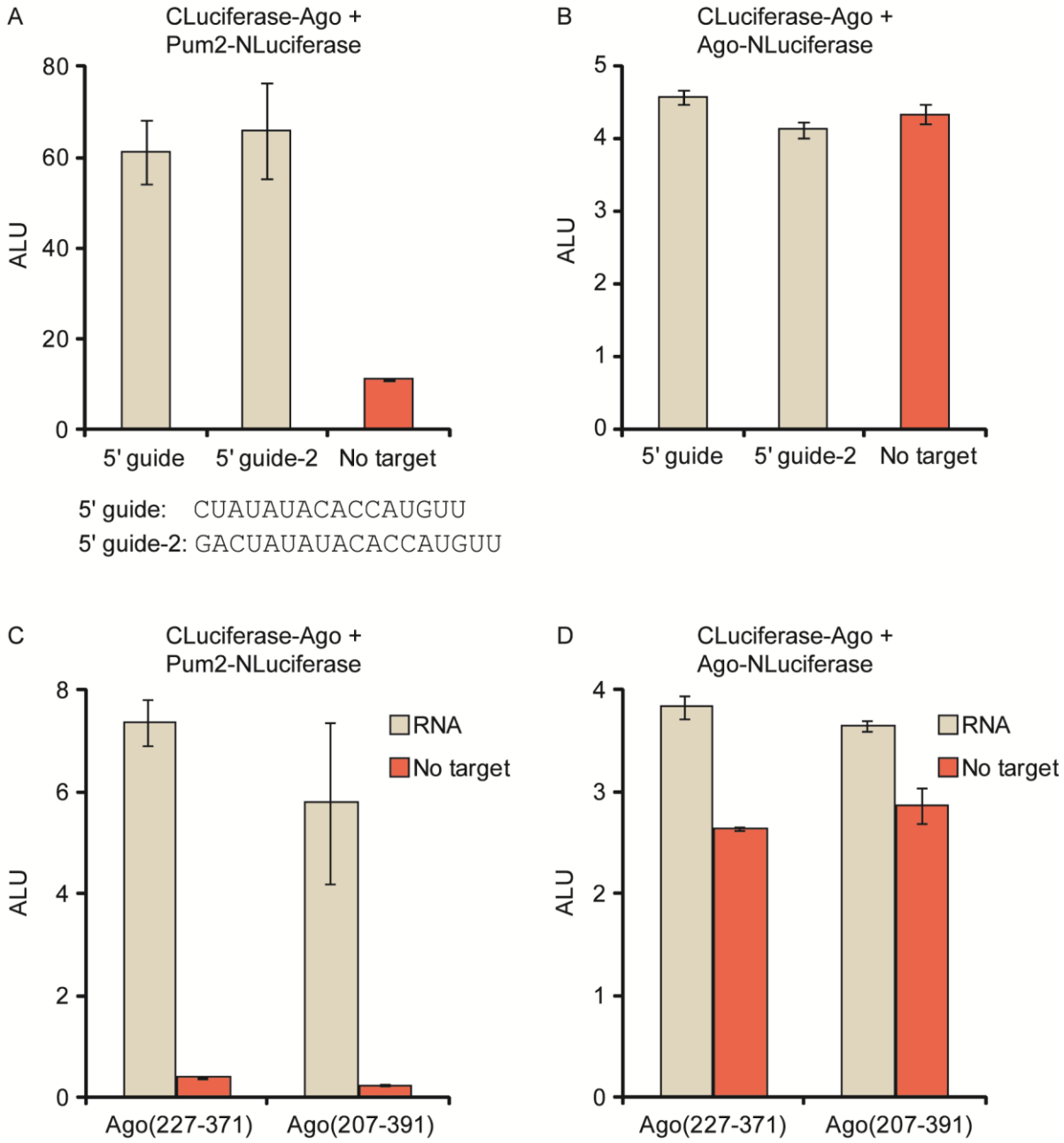

Figure S1. Optimization of the argonaute detection domain. (A) CLuciferase-Ago (Ago residues 227-371) and Pum2-NLuciferase were reassembled in the presence of the RNA target hybridized to either *5′ guide* or *5′ guide-2*, followed by luminescence readings (ALU). (B) CLuciferase-Ago and Ago-NLuciferase were reassembled in the presence of the RNA target hybridized to *3′ guide* and either *5′ guide* or *5′ guide-2*, followed by luminescence readings. (C) Two different Ago constructs, residues 207-391 or 227-371, were utilized as detection domains. CLuciferase-Ago(207-391) or CLuciferase-Ago(227-371) and Pum2-NLuciferase were reassembled in the presence of the RNA target hybridized to the *5′ guide*, followed by luminescence readings. (D) CLuciferase-Ago(207-391) and Ago(207-391)-NLuciferase or CLuciferase-Ago(227-371) and Ago(227-371)-NLuciferase were reassembled in the presence of the RNA target hybridized to the *3′ guide* and *5′ guide*, followed by luminescence readings.

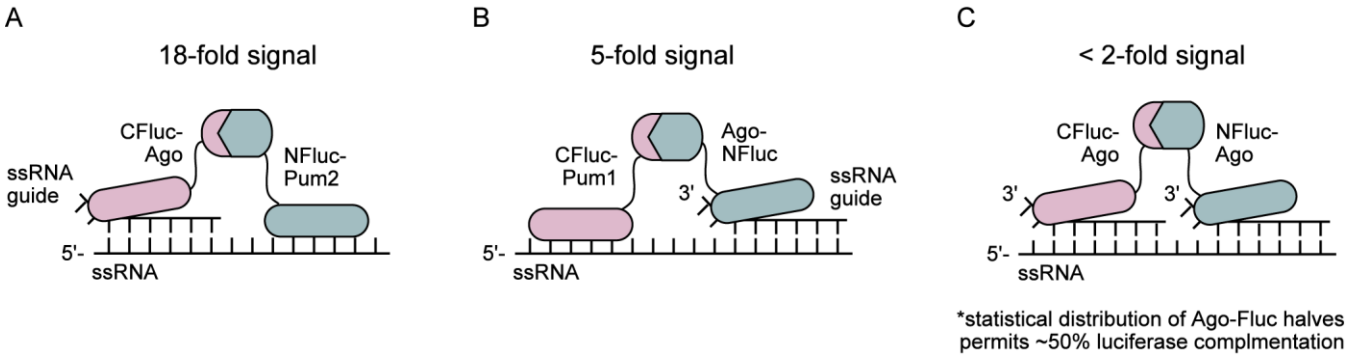

Figure S2. Limitations to RNA-templated split-luciferase reassembly using the Ago domain. (A) Ago can function effectively (18-fold signal) as a detection domain when a guide sequence is provided that is complementary to the 5' end of an RNA target. (B) When the 3' end of the guide sequence is present at an internal site in the target RNA, Ago affinity is reduced as judged by a reduced luminescence signal (4.5-fold signal). (C) Utilizing both guide sequences in tandem results in the expectation that a statistical distribution of Ago-luciferase halves bound to the RNA target would only permit ~50% split-luciferase complementation.

A

Length (AA)	Sequence
8	TGGGGSDI
14	TGGGGGSSGGGSDI
18	TGGGGGSGGGGSGGGSDI
23	TGGGGGSGGGGSGGGGSGGGSDI
28	TGGGGGSGGGGSGGGGSGGGGSGGGSDI

B

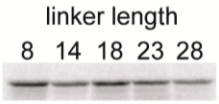

Figure S3. CLuciferase-linker-NLuciferase constructs. (A) The amino acid (AA) sequences of the linkers between CLuciferase and NLuciferase in the pcDNA-3.1(+)-CLuciferase-linker-NLuciferase constructs are presented. (B) Translations of CLuciferase-linker-NLuciferase constructs with 8, 14, 18, 23, and 28 AA linkers were performed in the presence of ^{35}S-methionine, and results were analyzed by SDS-PAGE.

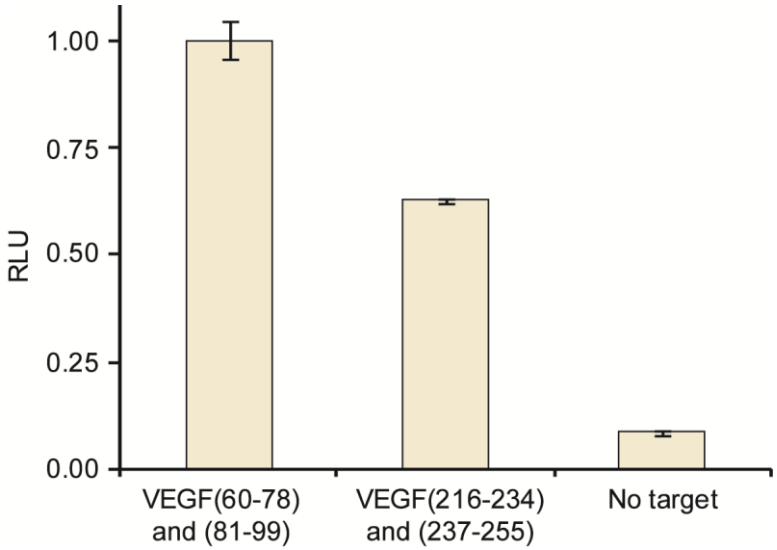

Figure S4. Generality of HP-guide design. HP-guides targeting regions 60-78/81-99 or 216-234/237-255 in the VEGF RNA template were annealed in the presence of the VEGF target. CLuciferase-E2C and Aart-NLuciferase sensors were incubated with the annealed products (1 nM), followed by luminescence readings. Results are presented as luminescence readings relative to the maximal signal (RLU).

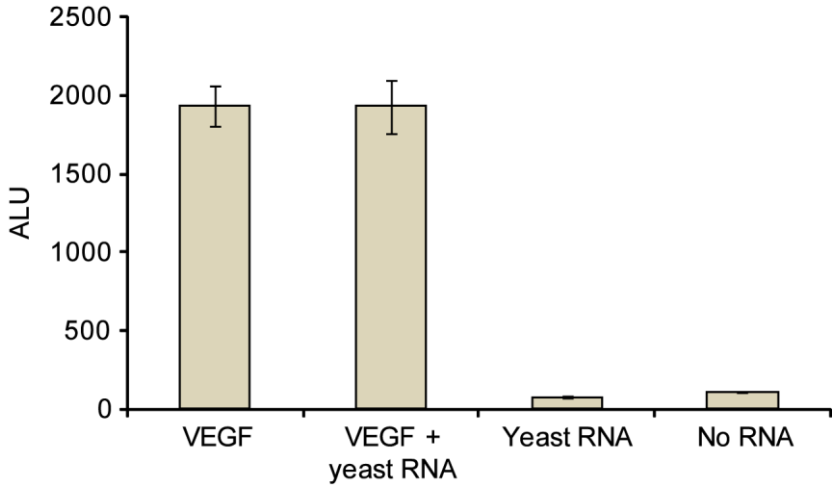

Figure S5. VEGF detection in the presence of yeast RNA. CLuciferase-E2C and Aart-NLuciferase were reassembled in the presence of 500 pM (5.0 ng) annealed VEGF target, 500 pM annealed VEGF + 10 μg yeast RNA (2000-fold mass excess), 10 μg yeast RNA, or No RNA. Luminescence readings (ALU) revealed that yeast RNA does not significantly affect VEGF RNA detection. The data presented in Figure 6F of the main text is a sub-set of the data presented here.

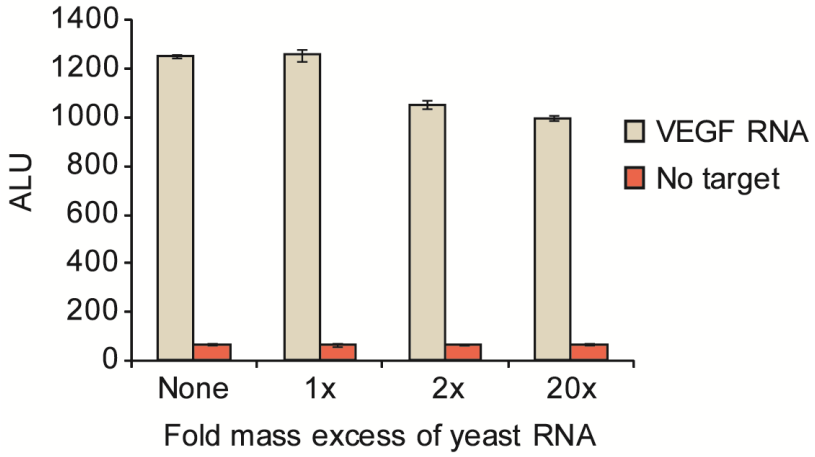

Figure S6. Effect of yeast RNA on VEGF annealing. VEGF was annealed to HP-guides in the presence of different mass excesses of yeast RNA. CLuciferase-E2C and Aart-NLuciferase were reassembled in the presence of the annealed VEGF target at 1 nM (10 ng) with no yeast RNA or with 10 ng (1x), 20 ng (2x), or 200 ng (20x) yeast RNA, followed by luminescence readings. The data indicate a slight decrease in annealing at 20x yeast RNA, although higher amounts were not tested due to experimental constraints on higher yeast concentrations.

Complete Reference 1 Citations

(1) (a) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. Q. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J. H.; Miklos, G. L. G.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, C.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z. M.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W. M.; Gong, F. C.; Gu, Z. P.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z. X.; Ketchum, K. A.; Lai, Z. W.; Lei, Y. D.; Li, Z. Y.; Li, J. Y.; Liang, Y.; Lin, X. Y.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B. X.; Sun, J. T.; Wang, Z. Y.; Wang, A. H.; Wang, X.; Wang, J.; Wei, M. H.; Wides, R.; Xiao, C. L.; Yan, C. H. *Science* **2001**, *291*, 1304-1351.

(b) Sachidanandam, R.; Weissman, D.; Schmidt, S. C.; Kakol, J. M.; Stein, L. D.; Marth, G.; Sherry, S.; Mullikin, J. C.; Mortimore, B. J.; Willey, D. L.; Hunt, S. E.; Cole, C. G.; Coggill, P. C.; Rice, C. M.; Ning, Z. M.; Rogers, J.; Bentley, D. R.; Kwok, P. Y.; Mardis, E. R.; Yeh, R. T.; Schultz, B.; Cook, L.; Davenport, R.; Dante, M.; Fulton, L.; Hillier, L.; Waterston, R. H.; McPherson, J. D.; Gilman, B.; Schaffner, S.; Van Etten, W. J.; Reich, D.; Higgins, J.; Daly, M. J.; Blumenstiel, B.; Baldwin, J.; Stange-Thomann, N. S.; Zody, M. C.; Linton, L.; Lander, E. S.; Altshuler, D. *Nature* **2001**, *409*, 928-933.

Supplementary References

(1) Meister, G.; Landthaler, M.; Patkaniowska, A.; Dorsett, Y.; Teng, G.; Tuschl, T. *Mol. Cell* **2004**, *15*, 185-197.
(2) Ma, J.-B.; Ye, K.; Patel, D. J. *Nature* **2004**, *429*, 318-322.
(3) Porter, J. R.; Stains, C. I.; Jester, B. W.; Ghosh, I. *J. Am. Chem. Soc.* **2008**, *130*, 6488-6497.
(4) Segal, D. J.; Crotty, J. W.; Bhakta, M. S.; Barbas III, C. F.; Horton, N. C. *J. Mol. Biol.* **2006**, *363*, 405-421.
(5) Beerli, R. R.; Segal, D. J.; Dreier, B.; Barbas III, C. F. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14628-14633.
(6) Cho, H.-S.; Mason, K.; Ramyar, K. X.; Stanley, A. M.; Gabelli, S. B.; Denney, D. W.; Leahy, D. J. *Nature* **2003**, *421*, 756-760.