Supporting Information

Click-RNA for Rapid Capture and Identification of Intracellular MicroRNA Targets

Pengjuan Zhang,[†,‡,§] Haohao Fu,[†,§] Shuangli Du,[†] Fengchao Wang,[†] Jie Yang,[†] Wensheng Cai,[*,†] and Dingbin Liu[*,†]

[†] *College of Chemistry, Research Center for Analytical Sciences, State Key Laboratory of Medicinal Chemical Biology, and Tianjin Key Laboratory of Molecular Recognition and Biosensing, Nankai University, Tianjin 300071 (China)*

[‡] *University of Environmental Engineering, Qinhuangdao, Hebei 066102 (China)*

Molecular simulations. The structures of PAZ and RNA were obtained from Protein Data Bank (www.rcsb.org, 4OLA, resid: 222-348) and generated by Avogadro (5'CUGAUGUUGA'3).[2] The initial structure of the protein-ligand hybrids in this study were prepared based on the binding site of PAZ identified in previous studies.[3-4] The protein-ligand complexes were then solvated in an equilibrated, cubic box of TIP3P water molecules.[5] The CHARMM 36m[6] and CHARMM general force field (CGenFF)[7] were utilized to describe biomolecules and DBCO, respectively.

All the atomistic MD simulations presented herein were performed using the parallel, scalable program NAMD 2.13.[8] Covalent bonds involving hydrogen atoms were constrained to their equilibrated lengths by means of the SHAKE/RATTLE[9-10] and SETTLE algorithms[11] for the organic and water molecules, respectively. Hydrogen-mass repartitioning[12] was used and a time step of 4 fs was applied to integrate the equation of motion. The particle mesh Ewald (PME) method[13] was employed to evaluate the long-range electrostatic forces. The temperature and the pressure were maintained at 300 K and 1 atm, respectively, using Langevin dynamics[14] and the Langevin piston method.[15] The short range van der Waals (VDW) and electrostatic interactions were calculated by the smoothed 12.0 Å spherical cutoff. Periodic boundary conditions (PBCs) were applied in the three directions of Cartesian space. Visualizations and analyses of the MD trajectories were performed with the VMD program.[16]

The MM-PBSA method[1] was employed to calculate the binding free energy of the PAZ:ligand complexes. The dielectric constants of the protein and the solvent were set to 2.0 and 78.54, respectively. The solvent radius was set to 1.4 A and the grid spacing was 0.5 A in each dimension. iAPBS,[17] the interface calling APBS[18] in NAMD, was used to calculate the Poisson–Boltzmann contribution to the binding free energy. Three 400 ns simulations characterizing PAZ, RNA and DBCO-tagged RNA in bulk water were performed to provide the trajectories needed in the free-energy calculations. The last 200-ns trajectories for protein-ligand complexes and the last 100-ns one were used for MM-PBSA calculation.

Supplementary Figures and Tables

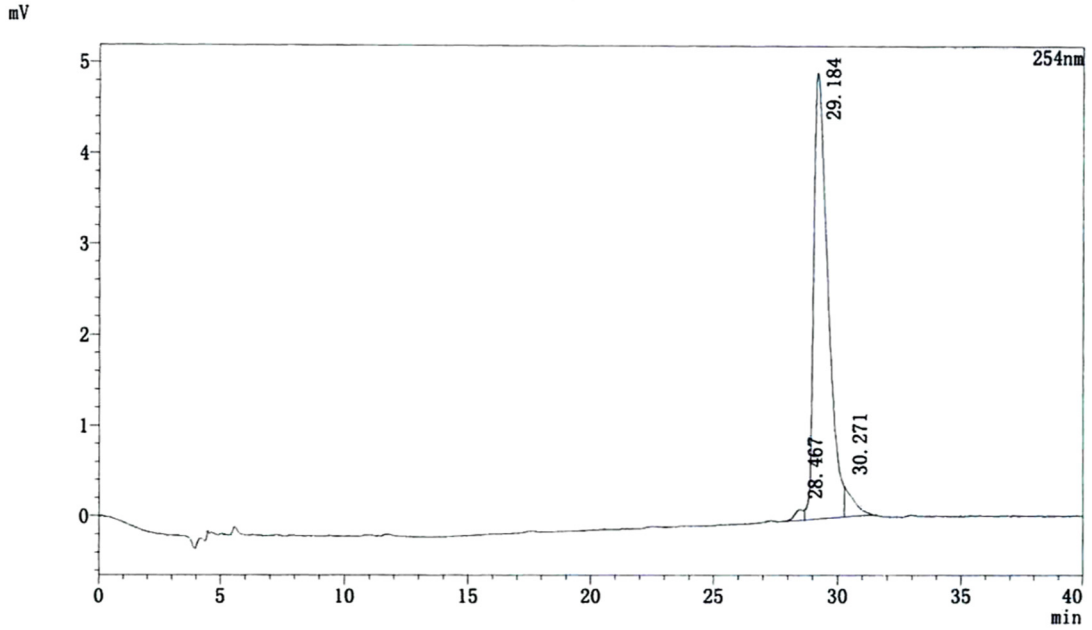

Figure S1. HPLC analysis of single-stranded 3'-NH_2-miR-21.

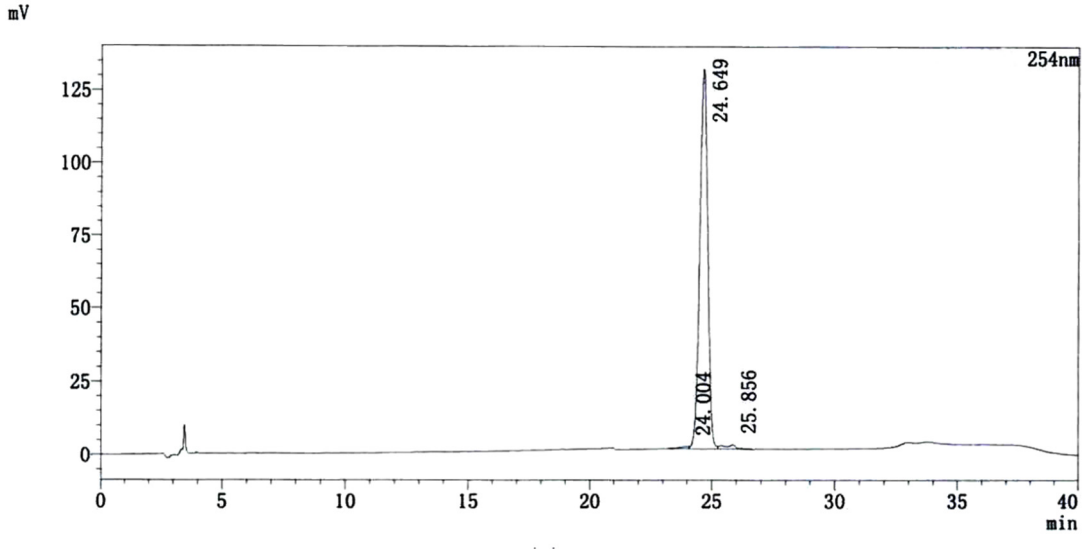

Figure S2. HPLC analysis of single-stranded 3'-DBCO-miR-21.

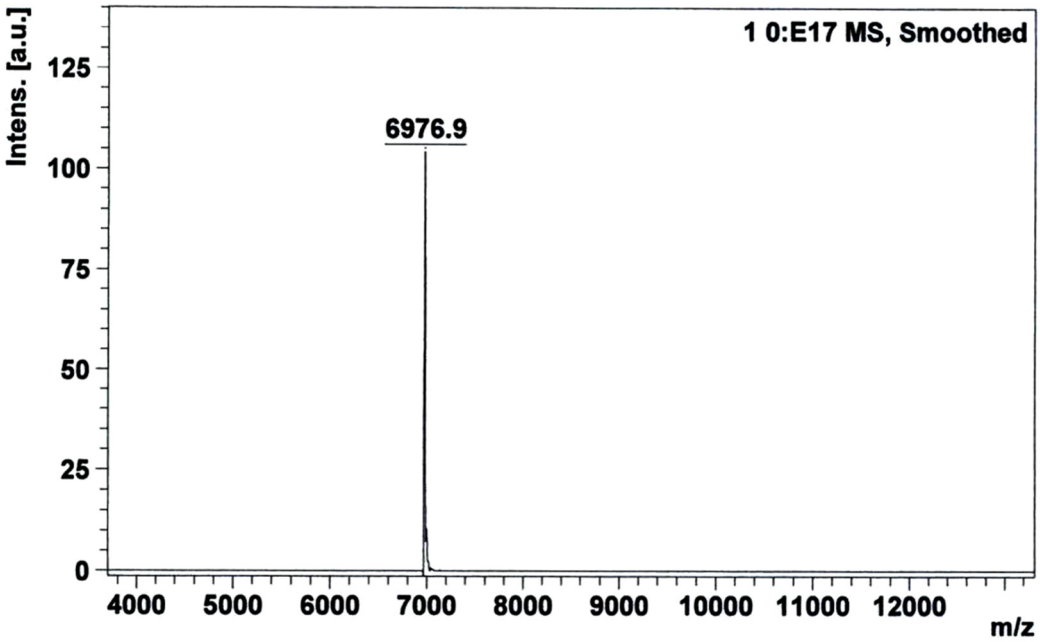

Figure S3. TOF-MS analysis of single-stranded 3'-NH_2-miR-21.

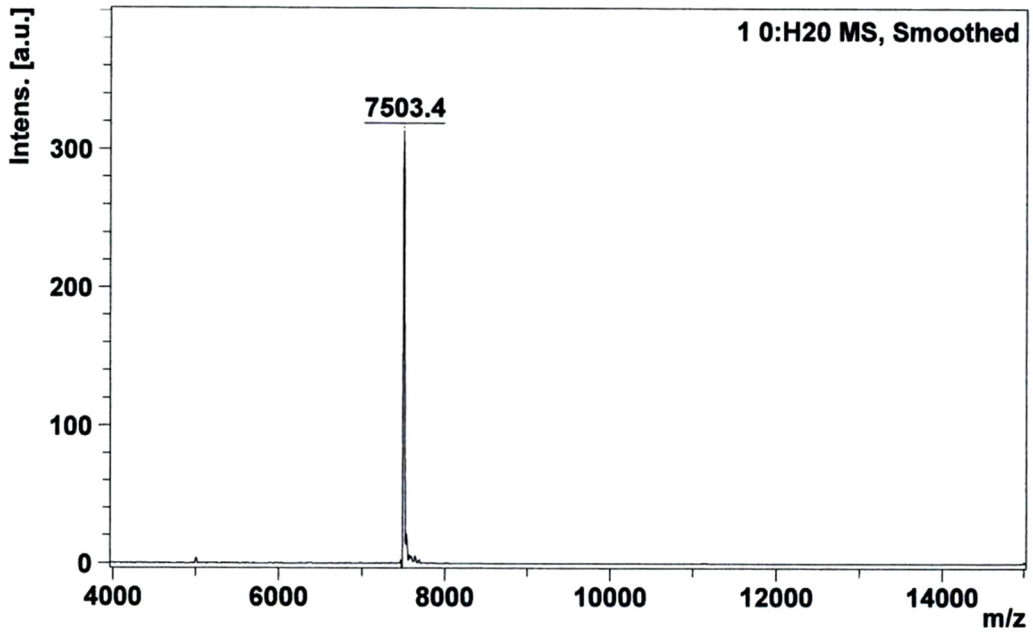

Figure S4. TOF-MS analysis of single-stranded 3'-DBCO-miR-21.

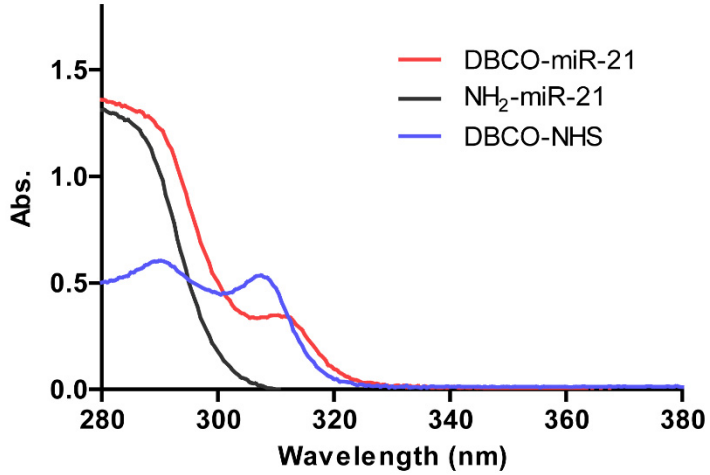

Figure S5. UV-spectrum characterization of double-stranded DBCO-miR-21. Compared with double-stranded NH_2-miR-21 (black), the specific peak of DBCO (blue) at around 310 nm could be observed for the DBCO-miR-21 (red).

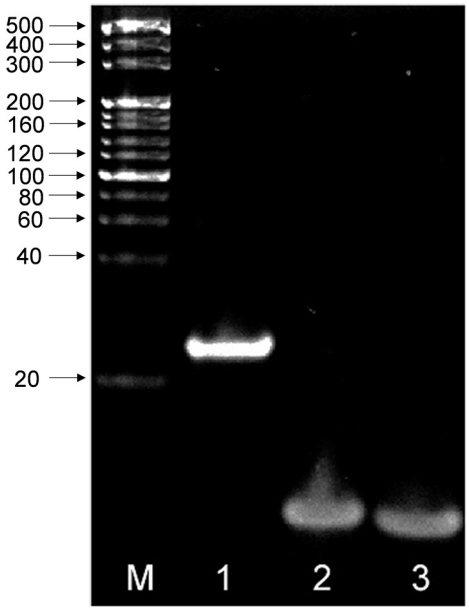

Figure S6. PAGE analysis of double-stranded DBCO-miR-21. M: 20bp Ladder marker; 1: the duplex of DBCO-miR-21 hybridized with its antisense strand; 2: single-stranded 3'-DBCO-miR-21; 3: the antisense strand of 3'-DBCO-miR-21.

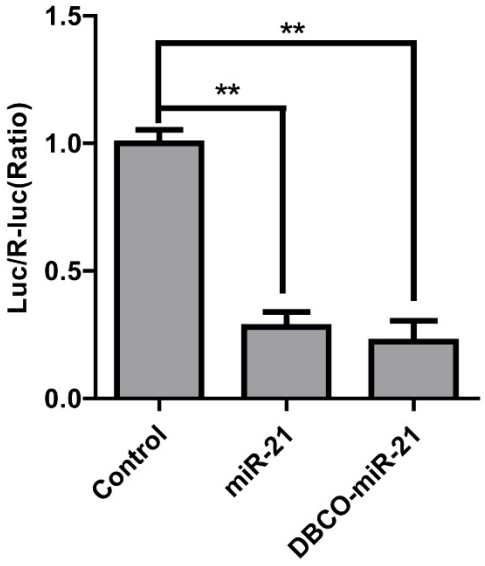

Figure S7. Function confirmation of DBCO-miR-21. HEK 293T cells were co-transfected pmiRGLO-PDCD4 with NC, miR-21 mimics, and DBCO-miR-21 respectively with the same concentration (30 nM). The luciferase activity was normalized to that of Renilla luciferase activity. Data are shown as mean ± SD (n = 6). **, $P < 0.01$.

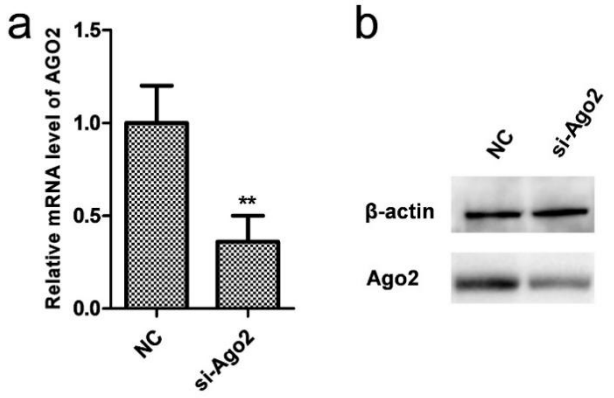

Figure S8. Relative mRNA expression levels and Ago2 protein abundance in HEK 293T cells after knockdown of Ago2. (a) qRT-PCR analysis of the mRNA expression levels of Ago2 in HEK 293T cells transfected with 30 nM of NC and si-Ago2, respectively. Data are shown as mean ± SD (n = 3). **, $P < 0.01$. (b) western blot analysis of Ago2 protein abundance in HEK 293T cells transfected with 30 nM of NC and si-Ago2, respectively.

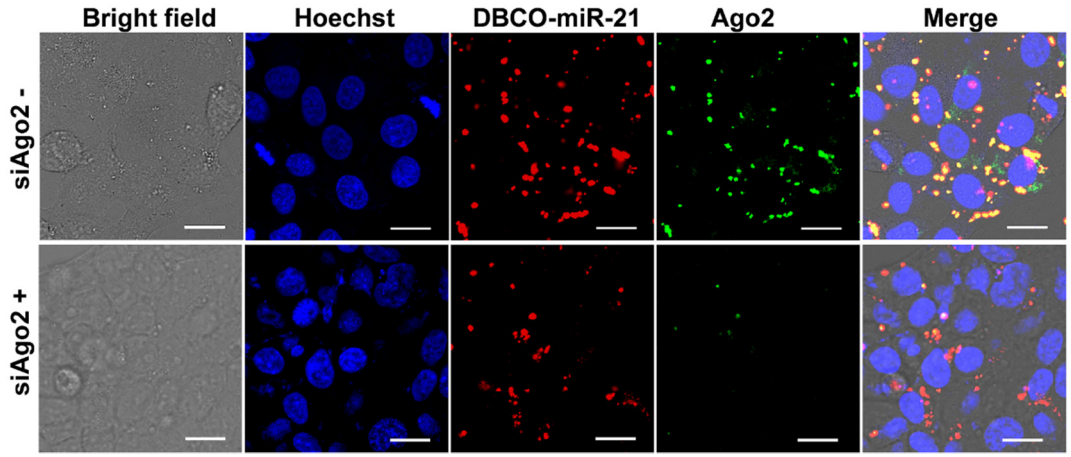

Figure S9. Co-localization of Ago2 and DBCO-miR-21 before and after knockdown of Ago2. HEK 293T cells were transfected with or without si-Ago2, and the cells were transfected with DBCO-miR-21 at 24 h post-transfection. The cells were then fixed, stained with azide-Cy5 (red) by click chemistry, immunofluorescence-stained with anti-Ago2 (green), counterstained with Hoechst (blue), and imaged with a Nikon A1+ Confocal Microscope. Scale bar: 20 μm.

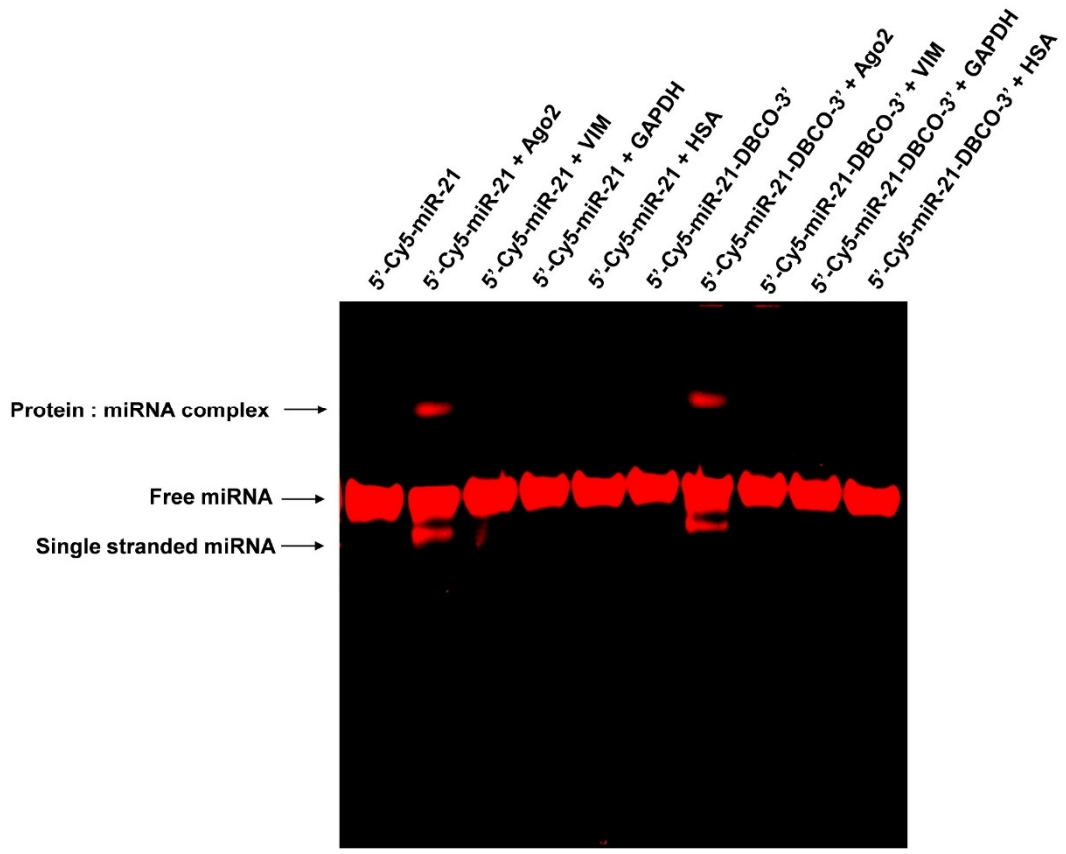

Figure S10. Binding specificity between DBCO-miR-21 and Ago2. Recombinant Ago2, VIM, GAPDH, and HSA (1 μg for each kind of proteins) were reacted with 10 nM of 5'-Cy5-miR-21-3'-DBCO mimics or 5'-Cy5-miR-21 mimics in the binding buffer (10 μL), respectively. Complexes were resolved by 5% native PAGE at 4 °C and imaged with Azure c600 (Azure Biosystems, USA).

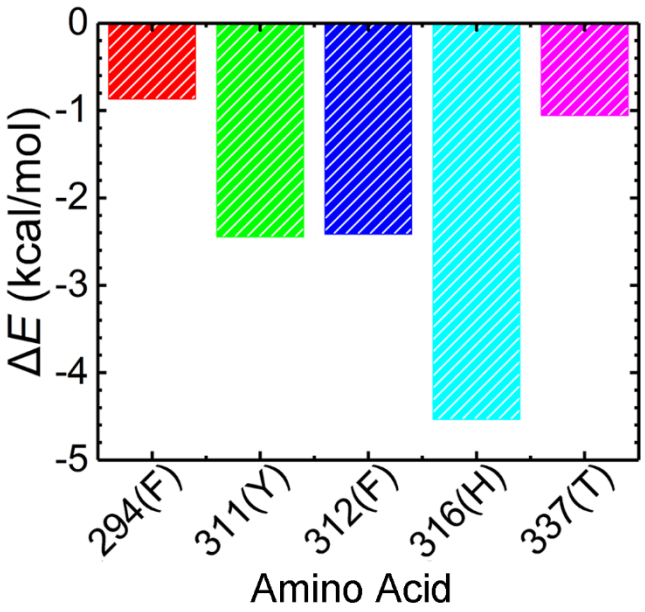

Figure S11. Interaction energy between DBCO and its adjacent amino acids.

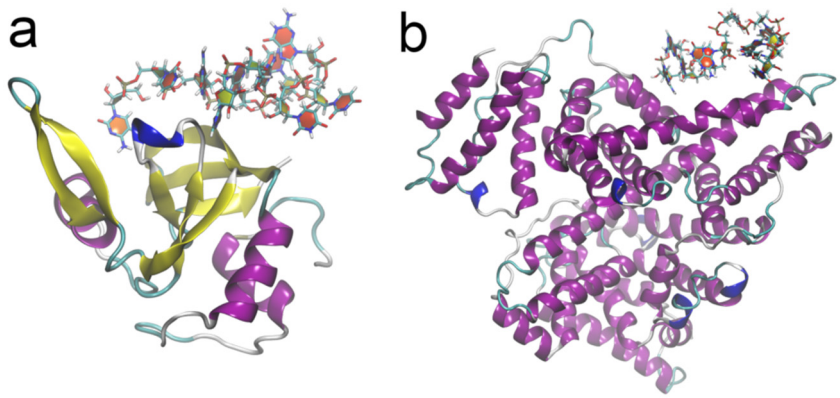

Figure S12. Three-dimensional arrangements of the (a) PAZ:DBCO-tagged miRNA (Adsorption motif) and (b) HSA:miRNA complex. The surface of HSA seems very hydrophilic such that the HSA:DBCO-tagged RNA hybrid is not energetically stable.

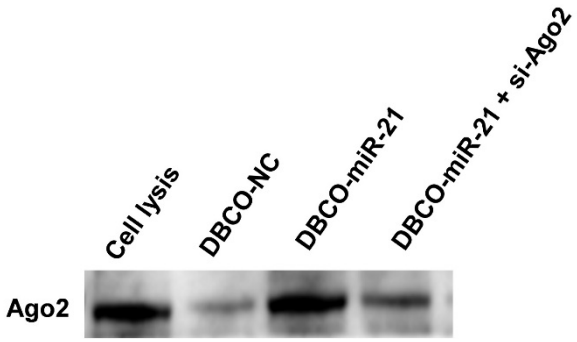

Figure S13. Western blotting analysis of Ago2 protein in the pulled-down lysate samples from the cells transfected with DBCO-miR-21 (100 nM).

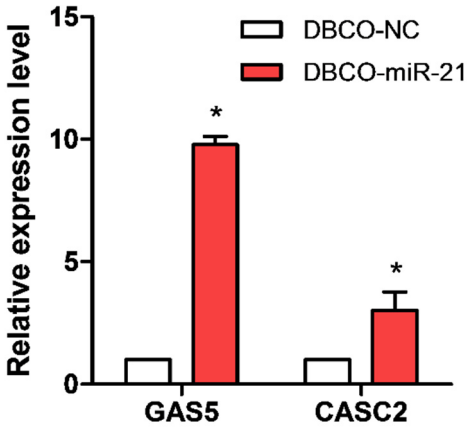

Figure S14. Application of DBCO-miR-21 as "bait" to pull down long non-coding RNA target genes inside cells. The target genes were identified with qRT-PCR analysis. Data are shown as mean ± SD (n = 3). *, $P < 0.05$.

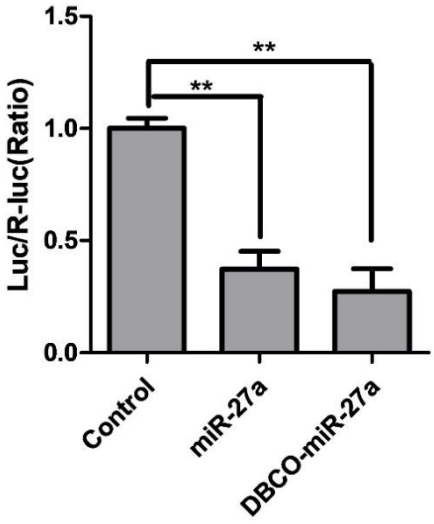

Figure S15. Function confirmation of DBCO-miR-27a. HEK 293T cells were co-transfected pmiRGLO-TGF-βRI with NC, miR-27a mimics, and DBCO-miR-27a respectively with the same concentration (30 nM). The luciferase activity was normalized to that of Renilla luciferase activity. Data are shown as mean ± SD (n = 6). **, $P < 0.01$.

Table S1. Absolute binding affinity (in kcal/mol) of DBCO-tagged miRNA towards Ago2 with respect to HSA.

Molecular assembly	Binding affinity
PAZ: DBCO-tagged miRNA (Insertion)	-539.0
PAZ: DBCO-tagged miRNA (Adsorption)	-439.3
HSA: DBCO-tagged miRNA	38.1

Table S2. Sequences of miR-21 and negative control (NC), and the 3'UTR sequences of PFKM, SKI, and TNKS.

miR-21	3'-AGUUGUAGUCAGAC**UAUUCGA**U-5'
NC	3'-GAUGAGAAAGAUCC**CCAACAC**U-5' (* * * * * *)
PFKM 3'UTR	274-5'···GAAUUUUCCUAA AA**AUAAGCU**U···3'-296
SKI 3'UTR	278-5'···UUUACACUUUUGUU**AUAAGCU**A···3'-300
TNKS 3'UTR	538-5'···CUGACAGCCUAGAA**AUAAGCU**G···3'-560

The bold font indicates the "seed" regions responsible for the recognition and binding between miR-21 and the target sequences. The asterisk represents the mismatches between NC and the miR-21 sequences in the "seed" region.

Table S3. Primers used for UTR clone, plasmid construction, and qRT-PCR.

Primers	Forward (5'-3')	Reverse (5'-3')	Application
BCL2	ATCGCCCTGTGGATGACTGAGT	GCCAGGAGAAATCAAACAGAGGC	qRT-PCR
c-Myc	CCTGGTGCTCCATGAGGAGAC	CAGACTCTGACCTTTTGCCAGG	qRT-PCR
WNT1	CTCTTCGGCAAGATCGTCAACC	CGATGGAACCTTCTGAGCAGGA	qRT-PCR
BTG2	GCAGAGGCTTAAGGTCTTCAGC	TGGTTGATGCGAATGCAGCGGT	qRT-PCR
TIMP3	TACCGAGGCTTCACCAAGATGC	CATCTTGCCATCATAGACGCGAC	qRT-PCR
NFIB	GGAACCAAGTCCTACAGGAGAC	GAATCCTGTGGAGATGCAGAGC	qRT-PCR
STAT3	CTTTGAGACCGAGGTGTATCACC	GGTCAGCATGTTGTACCACAGG	qRT-PCR
RECK	AATCCTTGCCCTGCCAATGAGC	GCACCTGGATTAGTGTCCCTTG	qRT-PCR
BMPR2	AGAGACCCAAGTTCCCAGAAGC	CCTTTCCTCAGCACACTGTGCA	qRT-PCR
PDCD4	ACTCCTAGAGCACCACAGTTGGT	CTCAGAAGCACGGTAGCCTTATC	qRT-PCR
PTEN	GAAGACCATAACCCACCACAGCT	ACACCAGTTCGTCCCTTTCCAG	qRT-PCR
TPM1	GAAGAGATCAAGGTCCTTTCCGAC	GGCTTTGTACTTCAGTTTCTGAGCG	qRT-PCR
TNKS	AGAGTACCTGCTACACCACGGT	AGTCCGCCACATTGACAGAAGC	qRT-PCR
SKI	CCTTCCGAAAAGGACAAGCCGT	GCTCTTTCTCACTCGCTGACAC	qRT-PCR
RBAK	AGCCGTGGATAATGGGAGGTGA	CTTTCTCTTCAGTCAGGGTTTTGC	qRT-PCR
PFKM	GCTTCTAGCTCATGTCAGACCC	CCAATCCTCACAGTGGAGCGAA	qRT-PCR
MCAM	ATCGCTGCTGAGTGAACCACAG	CTACTCTCTGCCTCACAGGTCA	qRT-PCR
ING3	AAGGCAACACGGTTCGCAGCTT	GCCATTTGCTCTTCCCTCCACT	qRT-PCR
BMP3	AAGGCAACACGGTTCGCAGCTT	GACTTGGTTAGCGATGTCAGATTG	qRT-PCR

ARMCX1	AGGAGTCTACGGACACCTCAGA	GTCCCAAACTCACCTCAGGCTT	qRT-PCR
DKK2	GGATGGCAGAATCTAGGAAGACC	CTGATGGAGCACTGGTTTGCAG	qRT-PCR
PDZD2	AGTGACTGTCGCTGGCTTTCAG	GACAGAGCACTGGCTAGTTCAC	qRT-PCR
GAPDH	CGAGCCACATCGCTCAGACA	GTGGTGAAGACGCCAGTGGA	qRT-PCR
GAS5	CTTCTGGGCTCAAGTGATCCT	TTGTGCCATGAGACTCCATCAG	qRT-PCR
CASC2	TACAGGACAGTCAGTGGTGGTA	ACATCTAGCTTAGGAATGTGGC	qRT-PCR
PDCD4	CGAGCTCGGACATTTTATAAACCTACAT	CCTGCAGGAATCAATACTGCTTCACATG	Plasmid construction
PFKM	CGAGCTCGCATGACTTCTGCCCCAGCTT	CCTGCAGGAAGCACAGTGACCAGTTGGC	Plasmid construction
SKI	CGAGCTCGCAACTACAGCTGGAGACGGG	CCTGCAGGCAGATGGCGTCATGTGTTGG	Plasmid construction
TNKS	CGAGCTCGACAGCCGCAGAGCAGAAG	CCTGCAGGAAGGGAACAGTAGCAGTTGAGT	Plasmid construction
FOXO1	CTACGAGTGGATGGTCAAGAGC	CCAGTTCCTTCATTCTGCACACG	qRT-PCR
SFRP1	CAATGCCACCGAAGCCTCCAAG	CAAACTCGCTGGCACAGAGATG	qRT-PCR
PINK1	GTGGACCATCTGGTTCAACAGG	GCAGCCAAAATCTGCGATCACC	qRT-PCR
SMAD5	CAGGAGTTTGCTCAGCTTCTGG	GGTGCTGGTTACATCCTGCCG	qRT-PCR
PPARγ	AGCCTGCGAAAGCCTTTTGGTG	GGCTTCACATTCAGCAAACCTGG	qRT-PCR
TGFBRI	GACAACGTCAGGTTCTGGCTCA	CCGCCACTTTCCTCTCCAAACT	qRT-PCR
PLK2	CAACAATGGTGCTCACATGAGCC	GGAGCATCTGTTGCTGGGAAAAC	qRT-PCR
CYP3A4	CCGAGTGGATTTCCTTCAGCTG	TGCTCGTGGTTTCATAGCCAGC	qRT-PCR
TGFBRI	CGAGCTCGGAGGTGGTAGCTA AAGAACA	CCTGCAGGTAATGACTGAAGGAAATGGA	Plasmid construction
miR-21 mimics	UAGCUUAUCAGACUGAUGUUGA	UCAACAUCAGUCUGAUAAGCUA	Probe
miR-27a mimics	UUCACAGUGGCUAAGUUCCGC	GCGGAACUUAGCCACUGUGAA	Probe

Negative control	UCACAACCCCUAGAAAGAG UAG	CUACUCUUUCUAGGAGGUU GUGA	Probe
DBCO-miR-21	UAGCUUAUCAGACUGAUGU UGA-DBCO	UCAACAUCAGUCUGAUAAGC UA	Probe
DBCO-miR-27a	UUCACAGUGGCUAAGUUCC GC-DBCO	GCGGAACUUAGCCACUGUG AA	Probe
DBCO-NC	UCACAACCCCUAGAAAGAG UAG-DBCO	CUACUCUUUCUAGGAGGUU GUGA	Probe
Cy5-miR-21	Cy5-UAGCUUAUCAGACUGAUGU UGA	UCAACAUCAGUCUGAUAAGC UA	Probe
Cy5-miR-21-DBCO	Cy5-UAGCUUAUCAGACUGAUGU UGA-DBCO	UCAACAUCAGUCUGAUAAGC UA	Probe

References

(1) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998,** *120* (37), 9401-9409.

(2) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **2012,** *4* (1), 17.

(3) Ma, J.-B.; Ye, K.; Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **2004,** *429*, 318.

(4) Song, J.-J.; Liu, J.; Tolia, N. H.; Schneiderman, J.; Smith, S. K.; Martienssen, R. A.; Hannon, G. J.; Joshua-Tor, L. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat. Struct. Biol.* **2003,** *10*, 1026.

(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

(6) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.;

Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017,** *14* (1), 71.

(7) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D., Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010,** *31* (4), 671-690.

(8) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005,** *26* (16), 1781-1802.

(9) Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983,** *52* (1), 24-34.

(10) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977,** *23* (3), 327-341.

(11) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992,** *13* (8), 952-962.

(12) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **2015,** *11* (4), 1864-1874.

(13) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

(14) Uhlenbeck, G. E.; Ornstein, L. S. On the theory of the Brownian motion. *Phys. Rev.* **1930,** *36* (5), 823.

(15) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **1995,** *103* (11), 4613-4621.

(16) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996,** *14* (1), 33-38.

(17) Konecny, R.; Baker, N. A.; McCammon, J. A. iAPBS: a programming interface to the adaptive Poisson–Boltzmann solver. *Comput. Sci. Discov.* **2012,** *5* (1), 015005.
(18) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98* (18), 10037-10041.