A general fragment-based approach to identify and optimize bioactive ligands targeting RNA

Blessy M. Suresh[a], Weichao Li[a], Peiyuan Zhang[a], Kye Won Wang[b], Ilyas Yildirim[b], Christopher G. Parker[a], and Matthew D. Disney[a,1]

[a]Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458; and [b]Department of Chemistry and Biochemistry, Florida Atlantic University, Jupiter, FL 33458

Edited by Sarah A. Woodson, Johns Hopkins University, Baltimore, MD, and accepted by Editorial Board Member Peter B. Moore November 9, 2020 (received for review June 13, 2020)

RNAs have important functions that are dictated by their structure. Indeed, small molecules that interact with RNA structures can perturb function, serving as chemical probes and lead medicines. Here we describe the development of a fragment-based approach to discover and optimize bioactive small molecules targeting RNA. We extended the target validation method chemical cross-linking and isolation by pull-down (Chem-CLIP) to identify and map the binding sites of low molecular weight fragments that engage RNA or Chem-CLIP fragment mapping (Chem-CLIP-Frag-Map). Using Chem-CLIP-Frag-Map, we identified several fragments that bind the precursor to oncogenic microRNA-21 (pre-miR-21). Assembly of these fragments provided a specific bioactive compound with improved potency that inhibits pre-miR-21 processing, reducing mature miR-21 levels. The compound exerted selective effects on the transcriptome and selectively mitigated a miR-21–associated invasive phenotype in triple-negative breast cancer cells. The Chem-CLIP-Frag-Map approach should prove general to expedite the identification and optimization of small molecules that bind RNA targets.

nucleic acids | RNA | fragment-based drug discovery | chemical biology | cancer

The Encyclopedia of DNA Elements (ENCODE) project has shown that 1 to 2% of the human genome encodes for protein; yet 75% is transcribed into RNA (1). Functional studies have elucidated that noncoding (nc) RNAs exhibit regulatory functions that can be subdivided into multiples classes (e.g., long noncoding [lnc] RNAs, micro [mi] RNAs, small nuclear [sn] RNAs), and their dysregulation is associated with a variety of diseases (2, 3). The functions of many ncRNAs are directly correlated to their structure (4, 5). Thus, although the development of chemical probes that bind to structured regions in RNAs is a considerable challenge, these probes have the potential to elucidate biological function.

Indeed, our lead identification strategy, named Inforna, has designed many bioactive compounds by identifying small molecules that bind structured regions (6–8). Inforna has also been used to identify RNA-binding modules that bind different sites in the same RNA that are then assembled into a single ligand, increasing both affinity and selectivity (9–12). However, to augment the identification of novel small molecules that bind RNA, we sought to develop a new and synergistic approach to assemble fragments onto RNA targets as well as optimize ligands that emerge from Inforna. Fortuitously, photoaffinity fragments allow for the efficient exploration of low molecular weight chemical space and, via covalent cross-linking, allow identification of fragments with modest affinity.

Compared with conventional ligand discovery screens that use libraries of higher molecular weight compounds, fragment-based ligand discovery (FBLD) uses libraries of lower molecular weight compounds to more efficiently explore chemical space (13–15). However, a major challenge of FBLD is the development of assays that detect low-affinity binding or short residence times of fragments (14, 16–20). Recent work has demonstrated that this obstacle can be overcome through the installation of photoaffinity groups onto small molecule fragments to capture and identify bound protein targets in cells directly via mass spectrometry-based proteomics (21, 22). Thus, these fully functionalized fragments (FFFs) can be used to provide binding site information and broadly assess proteome-wide "ligandability," and can even be advanced to compounds that selectively modulate protein function.

In the present work, we integrated FFFs with tools for studying transcriptome-wide RNA ligandability—specifically, chemical cross-linking and isolation by pull-down (Chem-CLIP), which enables the identification of RNA targets and corresponding binding sites (23, 24). This integrated approach, dubbed Chem-CLIP fragment mapping (Chem-CLIP-Frag-Map), uses libraries of FFFs (21) to identify small molecules that engage specific RNA targets and maps their binding sites using Chem-CLIP-Map (23). We applied this strategy to develop a potent and selective bioactive ligand targeting the precursor to microRNA-21 (pre-miR-21), a ncRNA linked to cancers (25) and other indications (26–28) (Fig. 1*A*).

Like all miRNAs, miR-21 is transcribed in a nucleus as a primary miRNA (pri-miR-21) that is processed by the nuclease Drosha to afford the pre-miRNA (pre-miR-21) (29, 30). The pre-miRNA is then translocated to the cytoplasm and processed by Dicer to liberate the mature, functional miRNA that translationally represses mRNAs by binding complementary regions in their 3′ untranslated regions (UTRs). Chem-CLIP-Frag-Map was used to define ligands that bind to the Dicer processing site

Significance

The development of selective bioactive compounds that target RNA structures is difficult. Here we report an approach to identify low molecular weight fragments that bind to a cancer-causing RNA. By determining where the fragments bind within an RNA target, they can be assembled to provide a high-affinity and potent inhibitor. This approach could be a general way to develop bioactive small molecules that target RNA, which previously was considered "undruggable." It was applied to provide a precision lead medicine against an RNA that contributes to cancer metastasis.

Author contributions: M.D.D. conceived of the ideas; B.M.S., C.G.P., and M.D.D. designed research; B.M.S., W.L., P.Z., K.W.W., and I.Y. performed research; C.G.P. contributed new reagents/analytic tools; B.M.S., W.L., P.Z., K.W.W., I.Y., and C.G.P. analyzed data; and B.M.S., K.W.W., I.Y., C.G.P., and M.D.D. wrote the paper.

Competing interest statement: M.D.D. is a founder of Expansion Therapeutics.

This article is a PNAS Direct Submission. S.A.W. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: DISNEY@scripps.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2012217117/-/DCSupplemental.

First published December 14, 2020.

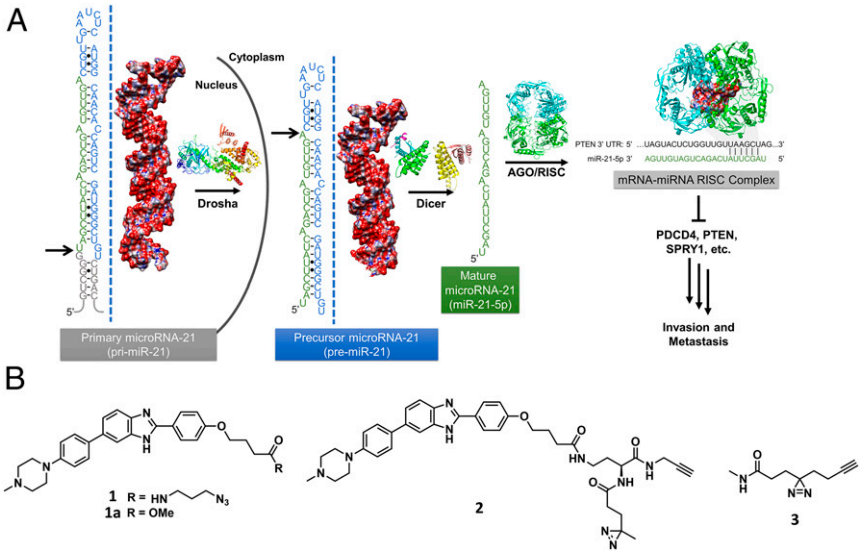

Fig. 1. The biogenesis of miR-21, its oncogenic function, and structures of lead molecules. (*A*) Scheme for miR-21 processing and its oncogenic function taken from ref. 9. Arrows in pri-miR-21 and pre-miR-21 point to Drosha and Dicer processing sites, respectively. The sequences and secondary structures of pri-miR-21, pre-miR-21, and mature miR-21 are from miRBase (55). The three-dimensional structures of the RNAs have been reported previously (9). (*B*) Chemical structures of **1** and its derivatives, **1a** and **2**, and control compound **3**. Compound **2** is a diazirine-functionalized version of **1** used to identify fragments that bind to pre-miR-21, while compound **3** is a control compound lacking the RNA-binding module. Compound **1a** was used in a competitive binding experiment to identify fragments that bind a different site than **1** in pre-miR-21 (Fig. 2).

of pre-miR-21 to selectively reduce the levels of mature miR-21 both in vitro and in cells.

Results and Discussion

Profiling Fragments that Bind RNA. Compound **1** (Fig. 1*B*) is a small molecule designed by Inforna that binds to the Dicer site in pre-miR-21 with a K_d value of 18 μM and modestly inhibits Dicer processing in the triple-negative breast cancer (TNBC) cell line MDA-MB-231 (9). We rationalized that compound **1** might be suitable for establishing a protocol to identify fragments that bind to RNA and also to read out their binding sites. To accomplish this goal, we synthesized compound **2,** which contains 1) an RNA-binding module based on compound **1**, 2) a diazirine group for the photoactivated capture of bound RNA targets (31–34), and 3) a terminal alkyne handle that can be bioorthogonally coupled to an azide-containing purification tag through click chemistry (35).

Compound **2** and the corresponding control compound **3**, which contains the diazirine and alkyne but lacks the RNA-binding module, were examined for their ability to bind and

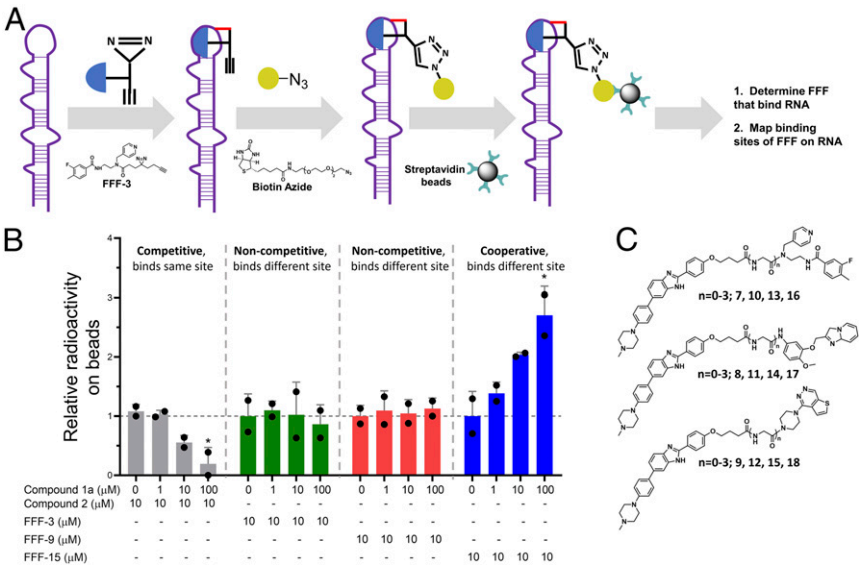

Fig. 2. A screening platform to identify binders of pre-miR-21 via Chem-CLIP and mapping of binding site via Chem-CLIP-Frag-Map. (*A*) Scheme for the Chem-CLIP assay using diazirine fragment reactivity to identify binders followed by mapping of the binding site. (*B*) The ability of each FFF to react with pre-miR-21 in presence of unreactive parent **1a** was measured. Compound **1a** (chemical structure in Fig. 1*B*) was unable to reduce the reaction of FFF-3, FFF-9, or FFF-15 with pre-miR-21, suggesting that they bind to different binding sites ($n = 2$). Error bars represent SD. *$P < 0.05$, one-way ANOVA compared with the reaction in which **1a** is absent. (*C*) Design of fragment assembly with varying linker lengths with glycine spacing modules.

cross-link to pre-miR-21 using the Chem-CLIP approach (*SI Appendix*, Fig. S1). A reaction containing 5′-^{32}P–labeled pre-miR-21 and **2** was irradiated with UV light and then subjected to a click reaction with biotin azide (biotin-N_3). The subsequent cross-linked adducts were pulled down and purified with streptavidin beads. We observed dose-dependent enrichment of pre-miR-21 with **2** but not with control compound **3** (*SI Appendix*, Fig. S1*A*). We further confirmed that coincubation with **1** competed with the cross-linking of **2** to pre-miR-21 in dose-dependent fashion (*SI Appendix*, Fig. S1*B*). As a secondary validation of this method, we fluorescently labeled the cross-linked RNA by substituting tetramethylrhodamine azide (TAMRA-N_3) for biotin-N_3. Analysis by fluorescence gel imaging confirmed the dose-dependent labeling of pre-miR-21, and coincubation with **1** reduced labeling of the RNA by **2** (*SI Appendix*, Fig. S1 *C* and *D*). Thus, pre-miR-21 labeling requires the RNA-binding module, and **2** binds the same site as the parent compound from which it was derived, as was previously observed with a similar Chem-CLIP probe in which the cross-linking moiety was an alkylating derivative (9).

Next, we determined whether we could map the binding site(s) of **2** using Chem-CLIP-Map (9, 23). In this approach, the RNA is cross-linked to the Chem-CLIP probe, followed by analysis by primer extension, or a reverse-transcription (RT) reaction. Reverse transcriptase is unable to proceed through the cross-linked sites, resulting in an "RT stop," the position of which can subsequently be determined from dideoxy sequencing. Indeed, three RT stops were observed upon cross-linking of **2** to pre-miR-21, at nucleotides G28, G37, and G38, near the A and U bulges that **1** bind (Fig. 1*B* and *SI Appendix*, Fig. S2). Collectively, this series of experiments verifies that diazirines cross-link to RNA and can be used to study compound binding and identify binding sites.

Fragment-Based Screening and Binding Site Mapping by Chem-CLIP-Frag-Map. As we have confirmed that diazirines coupled to RNA-binding ligands can map binding sites, we sought to determine whether this approach can be used to identify new RNA-binding ligands as well as map their sites of binding (Fig. 2*A*). A structurally diverse library of ∼450 FFF probes with photoactivable diazirine and alkyne moieties has been generated previously (21). Compounds in this FFF library were screened at 100 μM for binding of 5′-^{32}P–labeled pre-miR-21 in a 384-well format as described above in our validation studies. Following photocross-linking and conjugation of biotin-N_3, streptavidin-coated magnetic beads were used to separate and quantify the amount of cross-linked RNA. Of the 460 FFFs screened, 21 (a hit rate of 4.5%) cross-linked to pre-miR-21 to an extent exceeding three times the SD (3σ) from the mean for all compounds, a common metric for scoring hits in screening assays (36); 15 FFFs (indicated in red) showed greater enrichment of pre-miR-21 than **2** (*SI Appendix*, Fig. S3 *A* and *B*). Of the 21 hit FFF fragments, the majority had nitrogen-containing heterocycles, and 12 had nitrogen-containing aromatic heterocycles, structures commonly found in RNA binders (7).

A common strategy used to improve the affinity and potency of fragments for protein targets is to tether fragments together that engage neighboring binding sites (37, 38). Thus, we sought to demonstrate that this strategy can also be applied to RNA targets. We have previously shown that multivalency (i.e., tethering of RNA-binding modules) is a powerful approach for RNA targeting in general (10, 11, 39–41), and that a homodimer of **1** (**TGP-21**) targets the Dicer site in pre-miR-21 and inhibits its processing (9). Here we explored the tethering of low molecular weight fragments to **1**, each engaging a different structure within the RNA. A hybrid compound that emerges from such a strategy would provide ligands of lower molecular weight than **TGP-21** and might bind and recognize different features of the RNA's structure. **TGP-21** selectively recognizes the A and U bulges in pre-miR-21 (Fig. 1*A*) (9). An ideal fragment coupled to **1** would target the Dicer processing site and an adjacent structure but not necessarily the adjacent U bulge bound by the **1** dimer.

Therefore, we determined the binding sites of hit FFFs first through competitive (C-)Chem-CLIP to eliminate compounds that bind to the same site as **1**. Increasing reactivity of the FFF with the addition of **1** suggests that the two molecules bind cooperatively. Most of the FFFs (87%) either did not enrich RNA at the 10-fold lower concentration of 10 μM (vs. 100 μM in the

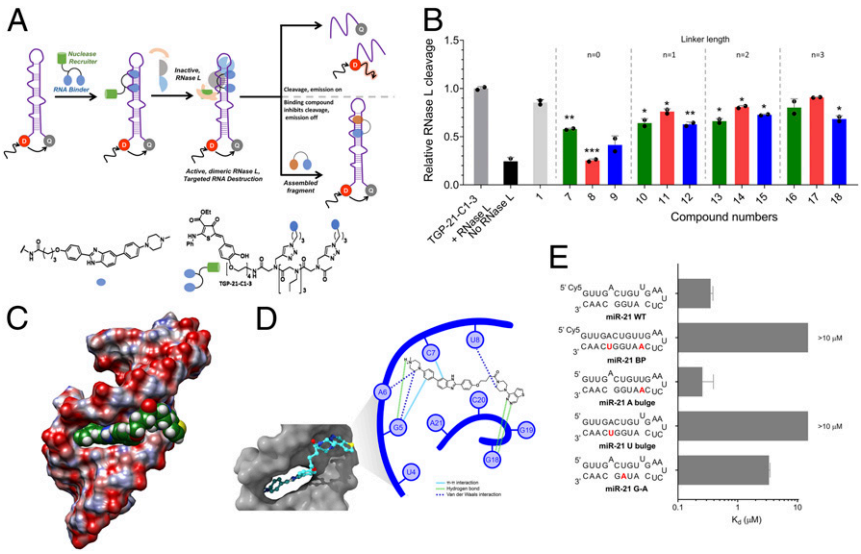

Fig. 3. RNase L cleavage assay to screen for binding of assembled fragments to pre-miR-21 and binding studies of hit compounds. (*A*) Scheme for the competitive RNase L cleavage assay. **TGP-21-C1-3** recruits RNase L to cleave the target RNA. Binding compounds compete with **TGP-21-C1-3** and prevent pre-miR-21 cleavage. (*B*) The assembled fragments protect pre-miR-21 from cleavage by RNase L. FAM fluorescence was measured 30 min after the addition of RNase L ($n = 2$); $^*P < 0.05$; $^{**}P < 0.01$ compared with **TGP-21-C1-3** alone, Student's *t* test. (*C*) A model of the structure of the pre-miR-21 Dicer site bound to **9** as determined by molecular modeling. (*D*) Interactions between **9** and the pre-miR-21 Dicer site. (*E*) Binding studies of **9** and pre-miR-21's Dicer site (miR-21 WT) and other mutated RNAs ($n = 2$). The affinity of **9** (K_d values) for miR-21 WT, miR-21 BP, miR-21 A bulge, miR-21 U bulge, and miR-21 G-A were 0.336 ± 0.1 μM, >10 μM, 0.23 ± 0.1 μM, >10 μM, and 3.3 ± 0.6 μM, respectively. Error bars represent SD for all panels.

primary screening assay) or had reduced reactivity in the presence of **1**, suggesting that the A and U bulges are particularly ligandable. In both instances, these fragments are not suitable for constructing a hybrid compound with **1** and thus were eliminated from further consideration. Two hits, **FFF-3** and **FFF-9**, showed no competition even upon addition of a 10-fold excess of **1**, while one hit, **FFF-15**, showed an increase in reactivity suggestive of cooperative binding (Fig. 2*B*).

Akin to the studies described to map the binding sites of **2**, we determined the binding sites of the three lead FFFs from competitive Chem-CLIP (C-Chem-CLIP) studies, **FFF-3**, **FFF-9**, and **FFF-15**. **FFF-3** and **FFF-9** were found to react with nucleotides G28 and C34, respectively, on opposite sides of the U bulge, and **FFF-15** was found to react primarily with nt G28, with minor reactivities also observed at nt U33, G38, and C39 (*SI Appendix*, Fig. S2). Thus, using Chem-CLIP-Frag-Map, we have shown that fragment-based probes can recognize different structural patterns in the RNA target.

Design of a Bivalent Ligand Library to Provide Improved Compounds Targeting pre-miR-21. We next used these molecular recognition fingerprints to design a bivalent RNA binder by tethering **FFF-3**, **FFF-9**, and **FFF-15** to Informa-derived **1**, which should garner enhanced selectivity and potency compared with **1** alone by binding two sites in the RNA simultaneously. Varying the distances between binding modules influences both the affinity and selectivity of modularly assembled ligands (40, 42, 43). Indeed, the ideal dimer positions the RNA-binding modules such that they mimic the distance between the RNA structures that they bind. Therefore, we synthesized three libraries of heterodimers in which the fragment **4** (derived from **FFF-3**), **5** (derived from **FFF-9**), or **6** (derived from **FFF-15**) (*SI Appendix*, Fig. S4) was separated from **1** by different distances, afforded by linkers composed of glycine amino acids (n = 0, 1, 2, 3) (Fig. 2*C*).

To identify the best dimer from the three series, we used a previously developed assay that uses a dually labeled pre-miR-21 and a small molecule that recruits an RNase that cleaves the RNA dubbed **TGP-21-C1-3**, a ribonuclease targeting chimera (RIBOTAC) (9). **TGP-21-C1-3** comprises a **1**-dimer (**TGP-21**) coupled to a heterocyclic compound that recruits RNase L (**C1-3**) to cleave pre-miR-21. The RIBOTAC selectively cleaves pre-miR-21 in cells and in an in vivo model of metastatic breast cancer (9).

In brief, pre-miR-21 was labeled with a 5′-fluorescein (FAM) and a 3′-black hole quencher. Binding of **TGP-21-C1-3** recruits RNase L to cleave the RNA, increasing the fluorescence signal. If a fragment-**1** dimer competes with the RIBOTAC, then the observed fluorescence signal will be reduced (Fig. 3*A*). To validate the assay, **1** and **TGP-21** (9) were studied, both of which protected pre-miR-21 from RNase L-mediated cleavage. Importantly, the IC_{50} for each compound in this assay is consistent with its binding affinity: **1** has an IC_{50} >50 μM and a K_d = 18 ± 4 μM, while **TGP-21** has an IC_{50} = 1.2 ± 0.8 μM and a K_d = 1 ± 0.1 μM (9) (*SI Appendix*, Fig. S5*A*).

We next assessed **4–1**, **5–1**, and **6–1** heterodimers at a single dose (25 μM) in this assay. The assembled fragments protected the RNA from cleavage to differing extents, and the most potent were carried forward to measure IC_{50} values (Fig. 3*B* and *SI Appendix*, Fig. S5*B*). Interestingly, the assembled fragments with the shortest linker lengths were the most potent compounds with IC_{50} values of 21 ± 3 μM for **7** (a **4–1** dimer), 12 ± 1 μM for **8** (a **5–1** dimer), and 13 ± 2 μM for **9** (a **6–1** dimer) (*SI Appendix*, Fig. S5 *C* and *D*). Thus, fragment assembly gives rise to novel binders, likely with increased binding affinity (studied below), as the IC_{50} for **1** in this assay is >50 μM.

Next, microscale thermophoresis was used to measure the binding affinity and selectivity of **7**, **8**, and **9** for pre-miR-21 (*SI Appendix*, Fig. S6*A*). All three compounds were more avid than **1**, with K_d values of 1,280 ± 50 nM (similar to **1**-dimer **TGP-21** and 18-fold more avid than **1**), 2,690 ± 120 nM, and 352 ± 30 nM (threefold more avid than **TGP-21** and 60-fold more avid than **1**), respectively. None of the compounds had measurable binding to a control RNA in which the ligand-binding site was mutated (*SI Appendix*, Fig. S6 *B* and *C*). Interestingly, the highest-affinity dimer, **9**, was derived from **FFF-15** whose reactivity with pre-miR-21 increased in the presence of **1** (Fig. 2*B*), bolstering the notion of positive cooperativity between the two RNA-binding modules. Furthermore, **9** binds with 60-fold greater affinity than **1**, with only a modest increase in molecular weight (673 g/mol for **9** vs. 552 g/mol for **1**).

To better understand the interactions leading to this increased affinity, a model of pre-miR-21 bound to **9** was generated using molecular dynamics simulations (Fig. 3*C*). In this model, the interaction is stabilized by **1** forming interactions with the A bulge (stacking and hydrogen bonding), while the fragment forms a bifurcated hydrogen bond with an adjacent G-U base pair, providing an explanation for the dramatic enhancements in affinity and selectivity.

We next performed mutational studies to validate the interactions identified from the molecular dynamics simulations (*SI Appendix*, Fig. S7*A*). Compound **9** retained its affinity upon mutation of the U bulge to a U-A base pair (i.e., miR-21 A bulge) (*SI Appendix*, Fig. S7*B*), whereas the affinity was abolished on mutation of the A bulge at the Dicer processing site to an A-U base pair (*SI Appendix*, Fig. S7*C*). When the G-U base pair, predicted to hydrogen-bond to the fragment, was mutated to A-U, the affinity of **9** decreased by 10-fold (*SI Appendix*, Fig. S7*D*), confirming the importance of these interactions for the assembled compound.

C-Chem-CLIP-Map was used to map the binding sites of **9** in vitro. Reactions at A35 and G38 of compound **2** (10 μM) were decreased on addition of **9** (10 μM) confirming that **9** binds to the predicted reactive sites of the A and U bulges (*SI Appendix*, Fig. S8 *A* and *B*). As expected, **9** inhibited Dicer processing of pre-miR-21 in vitro dose-dependently (*SI Appendix*, Fig. S9*A*). Importantly, this inhibition can be traced to the selective binding of the RNA as it was unable to inhibit the Dicer processing of a pre-miR-21 mutant in which the A and U bulges were mutated to base pairs (*SI Appendix*, Fig. S9*B*).

Cellular Assessment of the Effect Lead Optimized Compounds on miR-21 and Downstream Protein Targets. MDA-MB-231 TNBC cells highly express miR-21, causing an invasive phenotype (9, 44). This invasive character can be traced to the translational repression of the phosphatase and tensin homolog (PTEN) and programmed cell death 4 protein (PDCD4) by miR-21 (45–47). As **9** was both more avid and potent in vitro than **1**, **7**, and **8**, it was assessed for it ability to inhibit the biogenesis of miR-21 and subsequent deactivation of its downstream oncogenic pathway in MDA-MB-231 cells.

We first screened dimers **7**, **8**, and **9** for activity in MDA-MB-231 cells using a facile reporter system in which the 3′ UTR of PTEN is fused to luciferase. Notably, the 3′ UTR harbors a mutation such that it is responsive only to miR-21 levels (48) (*SI Appendix*, Fig. S10*A*). All three dimers increased luciferase activity to varying extents, indicating inhibition of miR-21, although the effect of **7** was not statistically significant at the concentrations tested (0.1, 1, and 10 μM) (Fig. 4*A*). A dose-dependent increase in luciferase activity was observed for **8** and **9**, with statistically significant increases observed at 10 μM of **8** and at all three concentrations of **9** tested (0.1, 1, and 10 μM). For comparison, treatment with 20 μM of **1** increased luciferase activity by approximately twofold, roughly the same increase observed with 10 μM of **8** (twofold more potent) and 1 μM of **9** (20-fold more potent) (Fig. 4*A*).

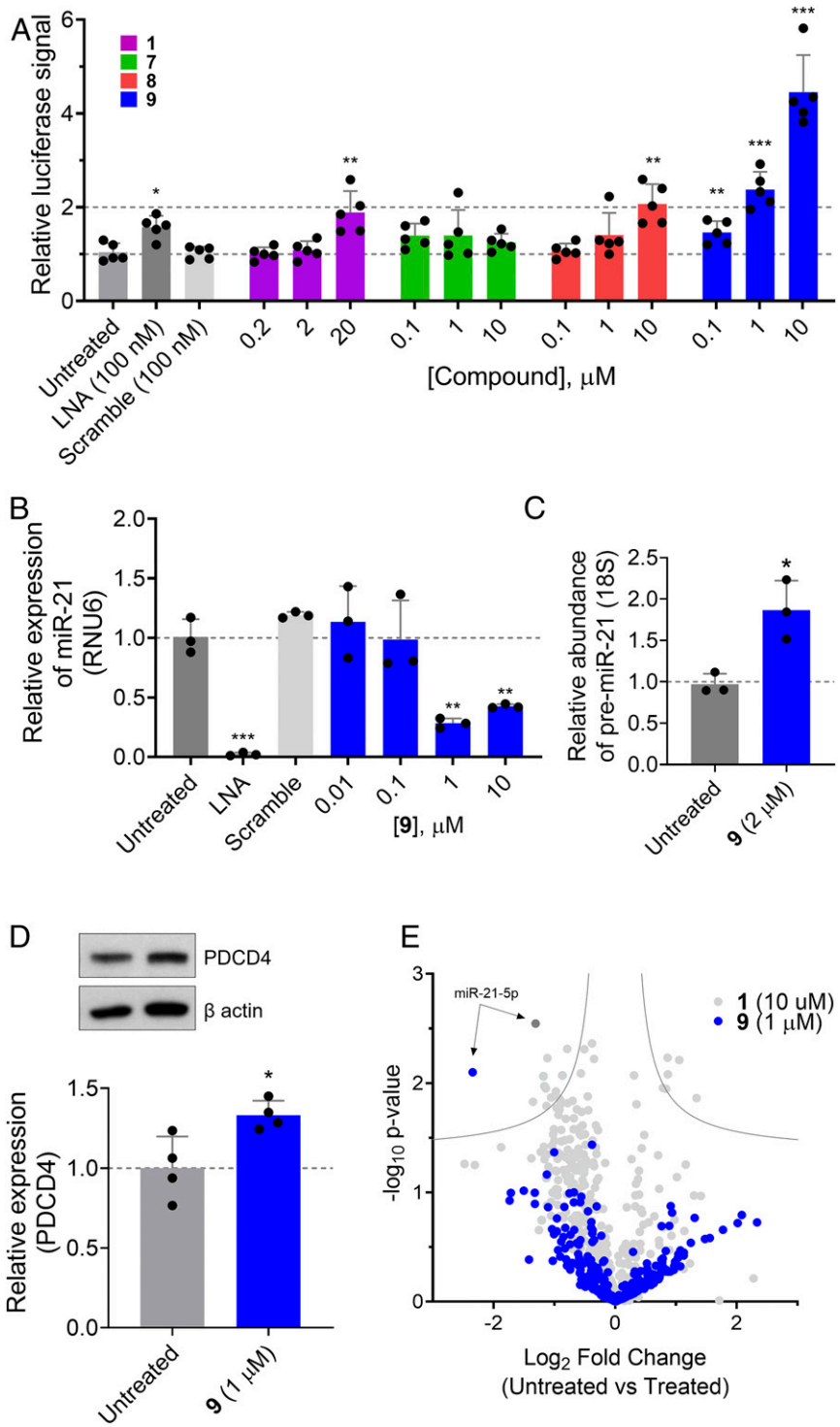

Fig. 4. Activity of fragment-based heterodimers in MDA-MB-231 TNBC cells. (*A*) Compounds **7**–**9** inhibit miR-21 functional activity, as assessed by a reporter assay in which the 3′ UTR of PTEN is fused to luciferase ($n = 5$). (*B*) Levels of mature miR-21 upon treatment with **9**, as measured by RT-qPCR ($n = 3$). (*C*) Levels of pre-miR-21 upon treatment with **9** (2 μM), as measured by RT-qPCR ($n = 3$). (*D*) Derepression of miR-21's downstream target PDCD4 upon treatment with **9** (1 μM), as measured by Western blot analysis ($n = 4$). (*E*) miRNA profiling of MDA-MB-231 cells treated with **9** (1 μM; $n = 3$). Dotted lines represent a false discovery rate (FDR) of 1% and variance of S0 (0.1). $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$, compared with untreated samples by a Student's *t* test. Error bars represent SD for all panels.

Considering the potency of **9** in this assay, we next assessed its ability to inhibit miR-21 biogenesis by measuring the levels of pre-miR-21 and mature miR-21 by RT-qPCR. If **9**'s mode of action is indeed inhibition of Dicer processing, then an increase in pre-miR-21 levels and reduction of mature miR-21 levels would be expected. Indeed, a dose-dependent decrease in mature miR-21 level was observed with an IC_{50} of ∼1 μM (Fig. 4*B*), and pre-miR-21 levels increased by ∼1.8-fold on treatment with

2 μM of **9** (Fig. 4*C*). Notably, **9** is 10-fold more potent than **1** and as potent as **TGP-21** (9), but with a lower molecular weight. [As a positive control for these studies, a locked nucleic acid (LNA) oligonucleotide, or antagomir, targeting miR-21 was used. As expected, the LNA (100 nM) reduced miR-21 levels and derepressed PTEN, while an LNA oligonucleotide with a scrambled sequence had no effect (Fig. 4).] This **9**-mediated inhibition of miR-21 biogenesis derepressed its endogenous target PDCD4 (46) by ∼40% at a dose of 1 μM (Fig. 4*D*), a 10-fold improvement over **1**, similar to the effects seen on mature miR-21 levels (9). Direct target engagement of pre-miR-21 by **9** was studied using C-Chem-CLIP. We first measured target engagement with **2** (1 μM), which provided a 2.8-fold enrichment in the amount of pre-miR-21; this enrichment was abolished upon the addition of **9** (1 μM), confirming direct target engagement (*SI Appendix*, Fig. S10*B*).

Compound 9 Is Selective miRNome-Wide. We next studied the selectivity of **9** (1 μM) across the miRNome by measuring its effect on the 379 miRNAs expressed in MDA-MB-231 cells. Only miR-21 levels were statistically significantly affected ($P < 0.01$) (Fig. 4*E*), whereas **1** (10 μM) affected the levels of 17 miRNAs (9). Thus, the addition of a fragment increased the potency and selectivity of an RNA binder in cells. Furthermore, these studies indicate that **9** is selective across the miRNome.

Fragment-Based Dimer 9 Reduces the Invasive Nature of MDA-MB-231 TNBC Cells. We next examined the effect of **9** on the miR-21–mediated invasive phenotype in MDA-MB-231 cells. Treatment of MDA-MB-231 cells with 1 μM of **9** (its approximate IC_{50} for inhibition of miR-21 biogenesis) reduced the number of invasive cells by ∼70% (Fig. 5 *A* and *B*). MCF-10A cells are a commonly used model of healthy human mammary epithelium that do not express miR-21 and thus have no invasive nature. Transfection of a plasmid expressing pre-miR-21 or a pre-miR-21 mutant (**9**-binding site ablated; *SI Appendix*, Fig. S5) caused MCF-10A cells to become invasive, as expected (Fig. 5 *C*–*E*). Importantly, the invasive nature of MCF-10A cells transfected with a plasmid encoding pre-miR-21, but not the mutant, was reduced on treatment with **9** (Fig. 5 *D* and *E*). Taken together, these experiments demonstrate that **9** has a specific effect on the invasive phenotype of TNBC by inhibiting miR-21 biogenesis.

Conclusion and Broader Impacts. In summary, we have described a fragment-based screening strategy to identify compounds that bind pre-miR-21 and map their sites of engagement, via Chem-CLIP-FragMap. Furthermore, we have provided a workflow for the assembly of fragments to develop high-affinity and selective bioactive compounds. The biological activity of the optimal designer compound **9** was demonstrated in MDA-MB-231, a TNBC cell line. Indeed, the fragment-based dimer was more potent and selective than the starting

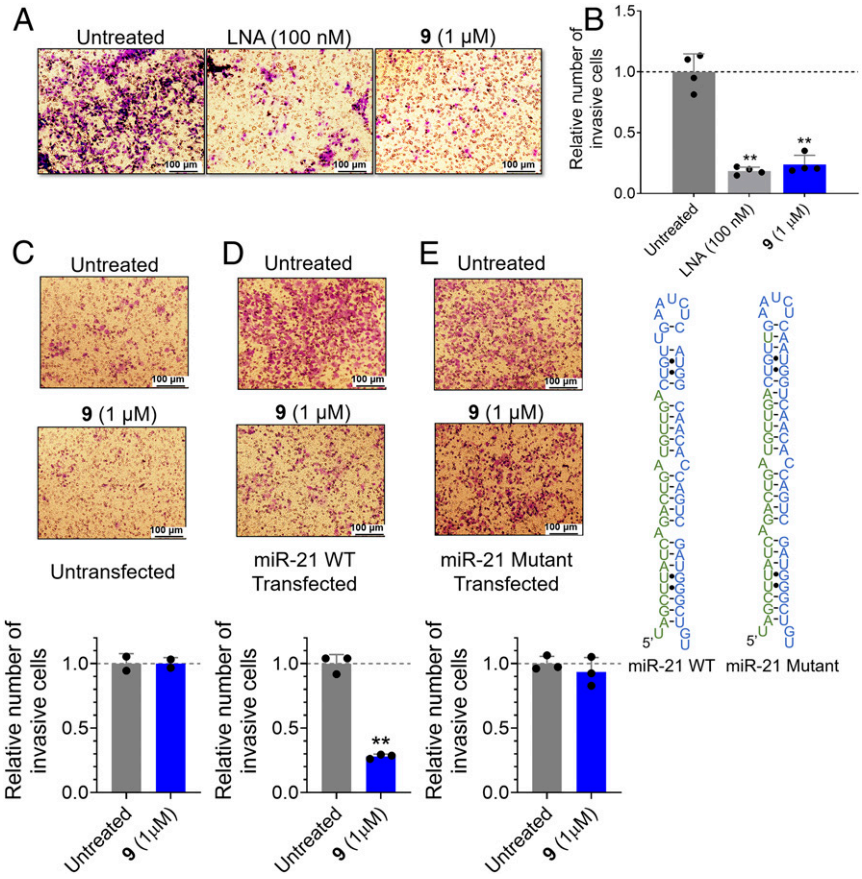

Fig. 5. Fragment-based dimer **9** inhibits miR-21–mediated invasion. (*A*) Representative images of MDA-MB-231 invasion assays on treatment with LNA targeting miR-21 or **9** (1 μM). (*B*) Quantification of the number of invasive MDA-MB-231 cells on treatment with LNA targeting miR-21 or **9** ($n = 4$). (*C*) MCF-10A cells, a model of healthy breast epithelium, do not expression miR-21 and thus have no invasive character or are affected on **9** treatment ($n = 2$). (*D*) Transfection of a plasmid encoding pre-miR-21 (miR-21 WT; secondary structure shown to the right) induces invasion in MCF-10A cells, which is abolished on **9** treatment ($n = 3$). (*E*) Transfection of a plasmid encoding a pre-miR-21 mutant in which the **9**-binding site is removed (miR-21 mutant; secondary structure shown at the right) also induces invasion in MCF-10A cells, but cannot be reversed by **9** treatment. $^{**}P < 0.01$ compared with untreated, as determined by a Student's *t* test. Error bars represent SD for all panels.

parent compound **1**, and, importantly, these enhancements were afforded by a very modest increase in molecular weight.

This ligand identification strategy likely can be widely applied to other disease relevant RNAs in vitro or in cells. Additional advancements in this area could include leveraging the foundation established here to study ligandability transcriptome-wide, followed by assembling bound fragments to afford potent and selective low molecular weight compounds. Despite the skepticism surrounding RNA as a drug target, these studies and others (49–54) suggest that chemical and chemical biology approaches can provide selective modulators of RNA (dys)function. Thus, it may be time to describe biomolecules that are perceived to be challenging small molecule targets as "not yet drugged" rather than "undruggable." As the science of chemical biology advances, it is becoming clear that more and more biomolecules are indeed targetable.

Data Availability. The atomic coordinates for compound **9** bound to pre-miR-21 have been deposited in ModelArchive at https://www.modelarchive.org/doi/10.5452/ma-izwdb (56). All other study data are provided in the main text and *SI Appendix*.

ACKNOWLEDGMENTS. We thank Jessica L. Childs-Disney for valuable help with manuscript preparation, Matthew Costales for experimental advice and helpful comments, and Jonathan Chen for assistance with creating figures for the RNA-ligand complex model structure. This work was supported by the NIH (R01 CA249180, to M.D.D.).

1. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
2. M. D. Disney, A. J. Angelbello, Rational design of small molecules targeting oncogenic noncoding RNAs from sequence. *Acc. Chem. Res.* **49**, 2698–2704 (2016).
3. R. Rupaimoole, F. J. Slack, MicroRNA therapeutics: Towards a new era for the management of cancer and other diseases. *Nat. Rev. Drug Discov.* **16**, 203–222 (2017).
4. Y. Wan *et al*., Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706–709 (2014).
5. Y. Ding *et al*., In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700 (2014).
6. S. P. Velagapudi, S. M. Gallo, M. D. Disney, Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat. Chem. Biol.* **10**, 291–297 (2014).
7. M. D. Disney *et al*., Inforna 2.0: A platform for the sequence-based design of small molecules targeting structured RNAs. *ACS Chem. Biol.* **11**, 1720–1728 (2016).
8. M. G. Costales *et al*., Small molecule inhibition of microRNA-210 reprograms an oncogenic hypoxic circuit. *J. Am. Chem. Soc.* **139**, 3446–3455 (2017).
9. M. G. Costales *et al*., Small-molecule targeted recruitment of a nuclease to cleave an oncogenic RNA in a mouse model of metastatic cancer. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2406–2411 (2020).
10. M. G. Costales *et al*., A designed small molecule inhibitor of a non-coding RNA sensitizes HER2 negative cancers to Herceptin. *J. Am. Chem. Soc.* **141**, 2960–2974 (2019).
11. S. P. Velagapudi *et al*., Design of a small molecule against an oncogenic noncoding RNA. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5898–5903 (2016).
12. A. J. Angelbello *et al*., Precise small-molecule cleavage of an r(CUG) repeat expansion in a myotonic dystrophy mouse model. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7799–7804 (2019).
13. S. D. Bembenek, B. A. Tounge, C. H. Reynolds, Ligand efficiency and fragment-based drug discovery. *Drug Discov. Today* **14**, 278–283 (2009).
14. D. E. Scott, A. G. Coyne, S. A. Hudson, C. Abell, Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **51**, 4990–5003 (2012).
15. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, Twenty years on: The impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605–619 (2016).
16. B. Lamoree, R. E. Hubbard, Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem.* **61**, 453–464 (2017).
17. D. A. Erlanson, B. J. Davis, W. Jahnke, Fragment-based drug discovery: Advancing fragments in the absence of crystal structures. *Cell Chem. Biol.* **26**, 9–15 (2019).
18. A. Pahl, H. Waldmann, K. Kumar, Exploring natural product fragments for drug and probe discovery. *Chimia (Aarau)* **71**, 653–660 (2017).
19. H. R. Nasiri *et al*., Targeting a c-MYC G-quadruplex DNA with a fragment library. *Chem. Commun. (Camb.)* **50**, 1704–1707 (2014).
20. P. N. Mortenson, D. A. Erlanson, I. J. P. de Esch, W. Jahnke, C. N. Johnson, Fragment-to-lead medicinal chemistry publications in 2017. *J. Med. Chem.* **62**, 3857–3872 (2019).
21. C. G. Parker *et al*., Ligand and target discovery by fragment-based screening in human cells. *Cell* **168**, 527–541.e29 (2017).
22. Y. Wang *et al*., Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nat. Chem.* **11**, 1113–1123 (2019).
23. S. P. Velagapudi, Y. Li, M. D. Disney, A cross-linking approach to map small molecule-RNA binding sites in cells. *Bioorg. Med. Chem. Lett.* **29**, 1532–1536 (2019).
24. L. Guan, M. D. Disney, Covalent small-molecule-RNA complex formation enables cellular profiling of small-molecule-RNA interactions. *Angew. Chem. Int. Ed. Engl.* **52**, 10010–10013 (2013).
25. A. Esquela-Kerscher, F. J. Slack, Oncomirs—MicroRNAs with a role in cancer. *Nat. Rev. Cancer* **6**, 259–269 (2006).
26. I. G. Gomez *et al*., Anti-microRNA-21 oligonucleotides prevent Alport nephropathy progression by stimulating metabolic pathways. *J. Clin. Invest.* **125**, 141–156 (2015).
27. L. E. Buscaglia, Y. Li, Apoptosis and the target genes of microRNA-21. *Chin. J. Cancer* **30**, 371–380 (2011).
28. T. Thum *et al*., MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **456**, 980–984 (2008).
29. D. P. Bartel, MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
30. D. P. Bartel, MicroRNAs: Target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
31. J. Wang, P. G. Schultz, K. A. Johnson, Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4604–E4612 (2018).
32. K. Nakamoto, Y. Akao, Y. Ueno, Diazirine-containing tag-free RNA probes for efficient RISC-loading and photoaffinity labeling of microRNA targets. *Bioorg. Med. Chem. Lett.* **28**, 2906–2909 (2018).
33. K. Nakamoto, K. Minami, Y. Akao, Y. Ueno, Labeling of target mRNAs using a photoreactive microRNA probe. *Chem. Commun. (Camb.)* **52**, 6720–6722 (2016).
34. K. Nakamoto, Y. Ueno, Diazirine-containing RNA photo-cross-linking probes for capturing microRNA targets. *J. Org. Chem.* **79**, 2463–2472 (2014).
35. H. C. Kolb, M. G. Finn, K. B. Sharpless, Click chemistry: Diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* **40**, 2004–2021 (2001).
36. F. Pukelsheim, The three sigma rule. *Am. Stat.* **48**, 88–91 (1994).
37. D. A. Erlanson, J. A. Wells, A. C. Braisted, Tethering: Fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
38. J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, K. M. Shokat, K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
39. S. G. Rzuczek *et al*., Precise small-molecule recognition of a toxic CUG RNA repeat expansion. *Nat. Chem. Biol.* **13**, 188–193 (2017).
40. S. G. Rzuczek *et al*., Features of modularly assembled compounds that impart bioactivity against an RNA target. *ACS Chem. Biol.* **8**, 2312–2321 (2013).
41. W. Y. Yang, R. Gao, M. Southern, P. S. Sarkar, M. D. Disney, Design of a bioactive small molecule that targets r(AUUCU) repeats in spinocerebellar ataxia 10. *Nat. Commun.* **7**, 11647 (2016).
42. J. L. Childs-Disney, P. B. Tsitovich, M. D. Disney, Using modularly assembled ligands to bind RNA internal loops separated by different distances. *ChemBioChem* **12**, 2143–2146 (2011).
43. J. L. Childs-Disney, J. Hoskins, S. G. Rzuczek, C. A. Thornton, M. D. Disney, Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol.* **7**, 856–862 (2012).
44. A. M. Krichevsky, G. Gabriely, miR-21: A small multi-faceted RNA. *J. Cell. Mol. Med.* **13**, 39–53 (2009).
45. N. M. McLoughlin, C. Mueller, T. N. Grossmann, The therapeutic potential of PTEN modulation: Targeting strategies from gene to protein. *Cell Chem. Biol.* **25**, 19–29 (2018).
46. L. B. Frankel *et al*., Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *J. Biol. Chem.* **283**, 1026–1033 (2008).
47. D. Iliopoulos, S. A. Jaeger, H. A. Hirsch, M. L. Bulyk, K. Struhl, STAT3 activation of miR-21 and miR-181b-1 via PTEN and CYLD are part of the epigenetic switch linking inflammation to cancer. *Mol. Cell* **39**, 493–506 (2010).
48. K. A. O'Donnell, E. A. Wentzel, K. I. Zeller, C. V. Dang, J. T. Mendell, c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843 (2005).
49. A. Donlic, A. E. Hargrove, Targeting RNA in mammalian systems with small molecules. *Wiley Interdiscip. Rev. RNA* **9**, e1477 (2018).
50. D. D. Vo *et al*., Targeting the production of oncogenic microRNAs with multimodal synthetic small molecules. *ACS Chem. Biol.* **9**, 711–721 (2014).
51. M. D. Shortridge, G. Varani, Structure-based approaches for targeting non-coding RNAs with small molecules. *Curr. Opin. Struct. Biol.* **30**, 79–88 (2015).
52. T. Hermann, Small molecules targeting viral RNA. *Wiley Interdiscip. Rev. RNA* **7**, 726–743 (2016).
53. A. C. Stelzer *et al*., Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.* **7**, 553–559 (2011).
54. C. M. Connelly, M. H. Moon, J. S. Schneekloth Jr, The emerging role of RNA as a therapeutic target for small molecules. *Cell Chem. Biol.* **23**, 1077–1090 (2016).
55. S. Griffiths-Jones, R. J. Grocock, S. van Dongen, A. Bateman, A. J. Enright, miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140–D144 (2006).
56. K. W. Wang *et al*, Assembled fragment compound targeting pre-miR-21. ModelArchive. https://www.modelarchive.org/doi/10.5452/ma-izwdb5. Deposited 21 October 2020.