Development of a Cell-Based Assay for the Detection of pre-miRNA-Protein Interactions

by

Sydney L. Rosenblum

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor Amanda L. Garner, Chair
Assistant Professor Kristin S. Koutmou
Associate Professor Patrick J. O'Brien
Professor Nils G. Walter

Sydney L. Rosenblum

slrosen@umich.edu

ORCID iD: 0000-0001-8231-6559

Dedication

I would like to dedicate this dissertation to my family, friends, and mentors who helped me along the path to where I am today.

Acknowledgements

I would first like to thank my advisor, Professor Amanda Garner, for her endless support and enthusiasm. I am immensely grateful for the trust she put in me throughout the years. I would also like to thank the members of my dissertation committee, Professor Kristin Koutmou, Professor Patrick O'Brien, and Professor Nils Walter, for their insight and advice. Thank you to my lab mates throughout the years. In particular, I would like to thank Arya Menon for the wonderful advice, both experimental and personal, for listening to my frequent ramblings, and encouraging me every step of the way. I am also grateful for the Program in Chemical Biology and larger University of Michigan community. Thank you as well to the custodial staff of NCRC Building 520.

I would like to acknowledge former Garner laboratory member Dan Lorenz for his invention of RiPCA and passing the project along to me. I would also like to acknowledge Steve Vander Roest of the University of Michigan Center for Chemical Genomics (CCG) for his invaluable help in executing HTS assays as well as Vladimir Simov and George Giambasu of Merck (MSD) for their collaboration on the RiPCA screening project.

I must also thank those who introduced me to this field of study. I would not be where I am today without Ms. Ann Moriarty, who was the first to introduce me to what a career in science could look like, and Professor Aaron Leconte, my undergraduate research advisor. Thank you, Aaron, for teaching me how to think like a scientist and laying the foundation for me to be successful in graduate school.

To my friends from all walks of life, thank you for adding so much joy to my life. I would like to especially thank those who went through graduate school with me: Anthony, Fabienne, Lara, Meredith, and Tyler. Going through this process was immeasurably more enjoyable with you all to turn to for science discussions and social relief.

Finally, to my family and loved ones, thank you for the unwavering support, encouragement, and love. I would like to especially thank my mom, dad, and sister. I could not have done this without you in my corner.

Table of Contents

List of Figures

List of Tables

Abstract

MicroRNA (miRNA), a class of small, non-coding RNA, are the product of a series of precise processing steps and responsible for regulating the translation of >60% of human protein-coding transcripts. Consequently, the dysregulation of miRNA levels has been linked to many human diseases, including cancers and neurodegenerative and cardiovascular diseases. In turn, RNA-binding proteins (RBPs) have been studied as regulators of miRNA biogenesis. Advancements in large-scale technologies have enabled the identification of proteins that bind specific sequences of precursor (pre-) miRNA and the discovery of disease-relevant miRNAs has inspired efforts to identify small molecule inhibitors of such interactions. To aid the study and identification of inhibitors of RPIs, *in vitro* and in cellular RPI detection systems have been developed. However, requirements of biochemical and cellular methods limit their utility, particularly for use with small, highly processed RNAs.

RiPCA, or RNA interaction with Protein-mediated Complementation Assay, an assay for the direct detection of RPIs in live cells, was developed to enable the validation and manipulation of pre-miRNA-protein interactions. In RiPCA, cells stably expressing the small subunit (SmBiT) of a split nanoluciferase (NanoLuc) fused to HaloTag are transiently co-transfected with a functionalized pre-miRNA probe and a plasmid encoding the RBP-of-interest fused to the large subunit (LgBiT) of NanoLuc. The pre-miRNA probe becomes covalently conjugated to SmBiT via HaloTag and subsequent interaction between the pre-miRNA and RBP drives the reconstitution of functional NanoLuc. Initially optimized using the RPI between the let-7 family of miRNA and the Lin28 RBPs, RiPCA was shown to detect the let-7/Lin28 interaction in both the cytoplasm and

the nucleus. Furthermore, RiPCA was capable of discerning sequence-specific binding preference of Lin28 as well as indicating the relative binding affinity of Lin28 and its individual RNA-binding domains.

These results encouraged the expansion of RiPCA for the detection of other functional pre-miRNA-protein interactions involving the RBPs hnRNP A1, Msi1, and Msi2. RiPCA was similarly capable of detecting the relative binding affinities of these RBPs for several pre-miRNA sequences, including let-7 family members and pre-miR-18a. In addition, the ability of RiPCA to detect site-specific binding was probed using a small library of pre-miRNA probes. While data reflected site-specific binding for Lin28, it was not shown for hnRNP A1, Msi1, and Msi2. Nevertheless, RiPCA demonstrated broad applicability of detecting pre-miRNA-RBP interactions.

Finally, to enable high throughput screening (HTS) of inhibitors of pre-let-7d/Lin28, RiPCA was miniaturized, and a partially automated workflow was optimized. A screen of ~18,000 small molecules derived from a curated library resulted in the identification of seven potential let-7/Lin28 inhibitors. Further characterization of the top hits is required to fully elucidate their mechanism of action and activity against let-7/Lin28 in cells.

Future efforts should focus on further engineering RiPCA to enable more precise detection as well as detection of RPIs involving other classes of RNAs, including mRNAs, lncRNAs, and expanded repeats. Overall, the technology reported herein promises to advance the characterization of RPIs and provide a platform for the discovery of RPI inhibitors.

CHAPTER 1

microRNA: Biogenesis, Function, Regulation, and Biological Significance

MicroRNAs (miRNAs), a class of small, non-coding RNA, are key players in the post-transcriptional regulation of over half of human protein-coding genes.[1] In their mature form, miRNAs are ~22 nucleotides in length and associate with Argonaute (AGO) proteins in the cytoplasm to carry out sequence-guided silencing of mRNA transcripts through translational repression and mRNA degradation.[2] Greater than 2,600 human miRNAs have been identified[3] and are involved in the regulation of virtually all processes, most notably developmental timing, cell proliferation and differentiation, and apoptosis.[4-5] The production, availability, and function of miRNAs is tightly regulated and, in addition to the proteins involved canonical miRNA processing, several RNA-binding proteins (RBPs) have been shown to selectively regulate the biogenesis, abundance, and function of miRNAs.[6-11] Accordingly, normal cellular function is dependent upon proper regulation of miRNA and their protein binding partners, and aberrant miRNA and RBP expression has been tied to several human diseases and cancers.[11-15]

1.1 MicroRNA Biogenesis

In the canonical biogenesis pathway (Fig. 1.1), miRNAs are derived from a long primary transcript, pri-miRNA, generated by RNA polymerase II that undergoes several processing steps.[16] First, the pri-miRNA is cleaved at the base of the characteristic three spiral turn hairpin by the microprocessor complex, composed of Drosha, an RNAse III enzyme, and DGCR8/Pasha, a

double-stranded RNA-binding domain (dsRBD) protein, to form the ~60-80 nucleotide (nt) pre-miRNA species containing a 2 nt 3' overhang.[17-19] The characteristic overhang allows pre-miRNAs to be exported to the cytoplasm by an exportin 5/RanGTP complex and further processed by Dicer, an RNAse III endonuclease. Dicer cleaves the terminal loop, producing a ~21-23 nt duplex. The dsRBP TRBP, which associates with Dicer, acts as a bridge, helping to load one strand of the duplex into a protein from the AGO family, forming the miRNA-induced silencing complex (miRISC).[20-21]

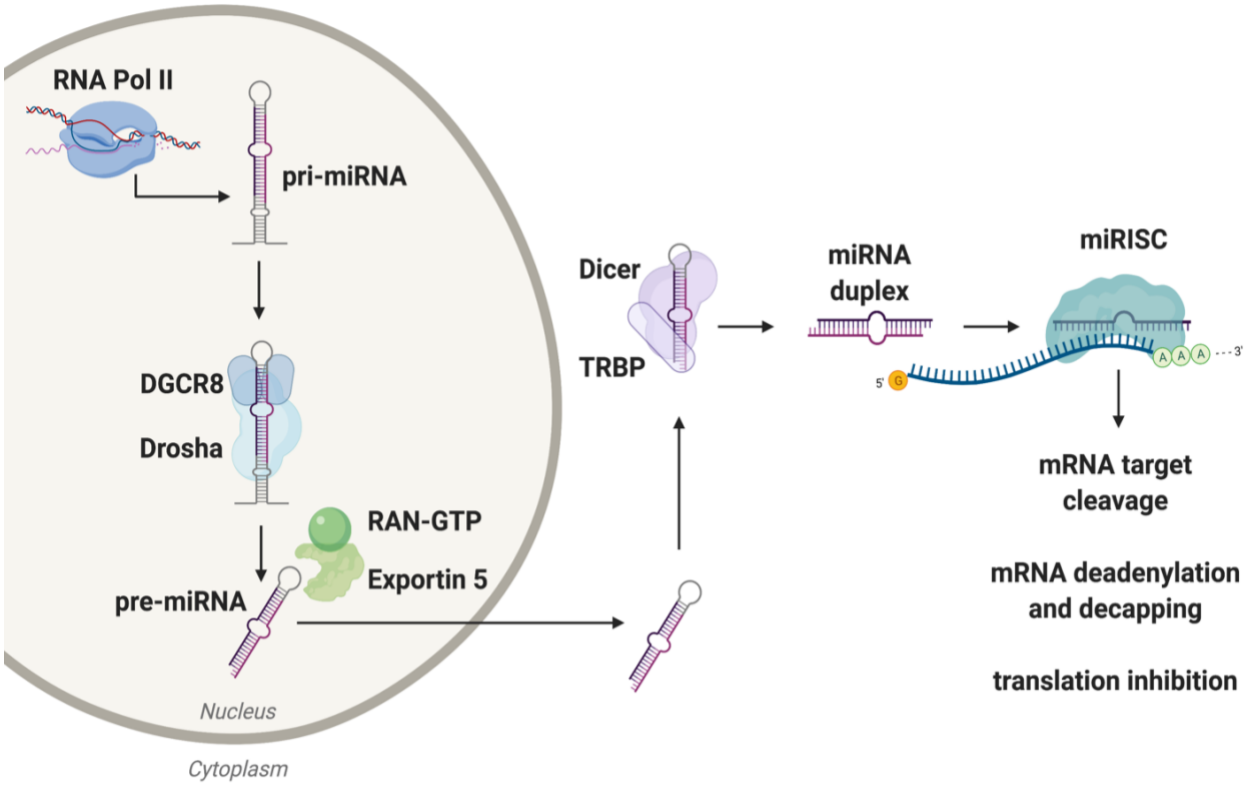


Figure 1.1. Canonical pathway of miRNA biogenesis. Created with BioRender.com.

1.2 Gene Regulation by miRNAs

Once loaded into miRISC, the miRNA guides the complex in a sequence-specific manner to the miRNA response element (MRE) of target mRNAs. Translational repression is then carried

out via two main mechanisms depending on the specific AGO protein that is contained in the miRISC and the complementarity between the miRNA sequence and the target mRNA.[22] Humans express four AGO proteins (1-4), AGO2 being the most abundant and the only one that possesses endonuclease activity.[23] Typically, miRNA binding sites are located within the 3' UTR, but have also been found in the 5' UTR and coding regions[24-25] of a target mRNA. If there is a central mismatch in the miRNA/MRE interaction, as is most common[26], gene silencing occurs through translational inhibition by disruption of the eukaryotic translation initiation factor 4F (eIF4F) complex formation, recruitment of poly(A)-deadenylases, and subsequent removal of the m^7G cap by the decapping complex. The uncapped and deadenlylated mRNA is destabilized and frequently degraded by an exonuclease. However, full complementarity between the miRNA and MRE induces AGO2-dependent endonucleolytic cleavage of the target mRNAs.[27]

1.3 Regulation of miRNA Levels

As expected, there is tight regulation of miRNA biogenesis at each step of the pathway.[2] The rate of transcription and processing efficiency of miRNA precursors are the main determinants of the level of a mature miRNA.[28] Transcriptional regulation of pri-miRNA transcripts can occur epigenetically through modifications to histones or DNA directly or through binding of sequence-specific transcription factors. Several studies have shown that hypermethylation of a miRNA encoding gene, a phenomenon that is often found in cancers, reduces miRNA levels.[19] Transcription factors can either promote or repress transcription of specific pr-miRNA transcripts. For example, the transcription factor TP53 promotes the expression of the miR-34a family of miRNA, whereas the transcription factors ZEB1 and ZEB2 repress expression of miR-200.[29-32] Post-transcriptional processing and localization of pri-miRNA transcripts have also been found to influence processing efficiency.[33] Additionally, post-translational modifications (PTMs) of the

proteins involved in miRNA biogenesis can alter their activity or specificity, impacting which miRNA sequences mature through the pathway.[34] Other factors, including the structure and sequence of miRNA precursors, can also affect the rate at which miRNAs are processed.[35-38]

After the serendipitous discovery of several auxiliary RBPs that affect miRNA processing, there has been increasing interest in mapping the role of auxiliary RBPs in the regulation of miRNA biogenesis and function. Through their sequence specific binding, RBPs influence miRNA levels by either enhancing or inhibiting the processing of miRNA precursors.[12, 39-41] The development of methods for the genome-wide interrogation of RPIs, including CLIP- and proteomics-based technologies, has led to the discovery of countless proteins that interact with miRNA at the various stages of processing (Fig. 1.2).[28, 42] Figure 1.2 features some prominent examples of RBPs known to regulate miRNA biogenesis, some of which will be discussed in further detail below.

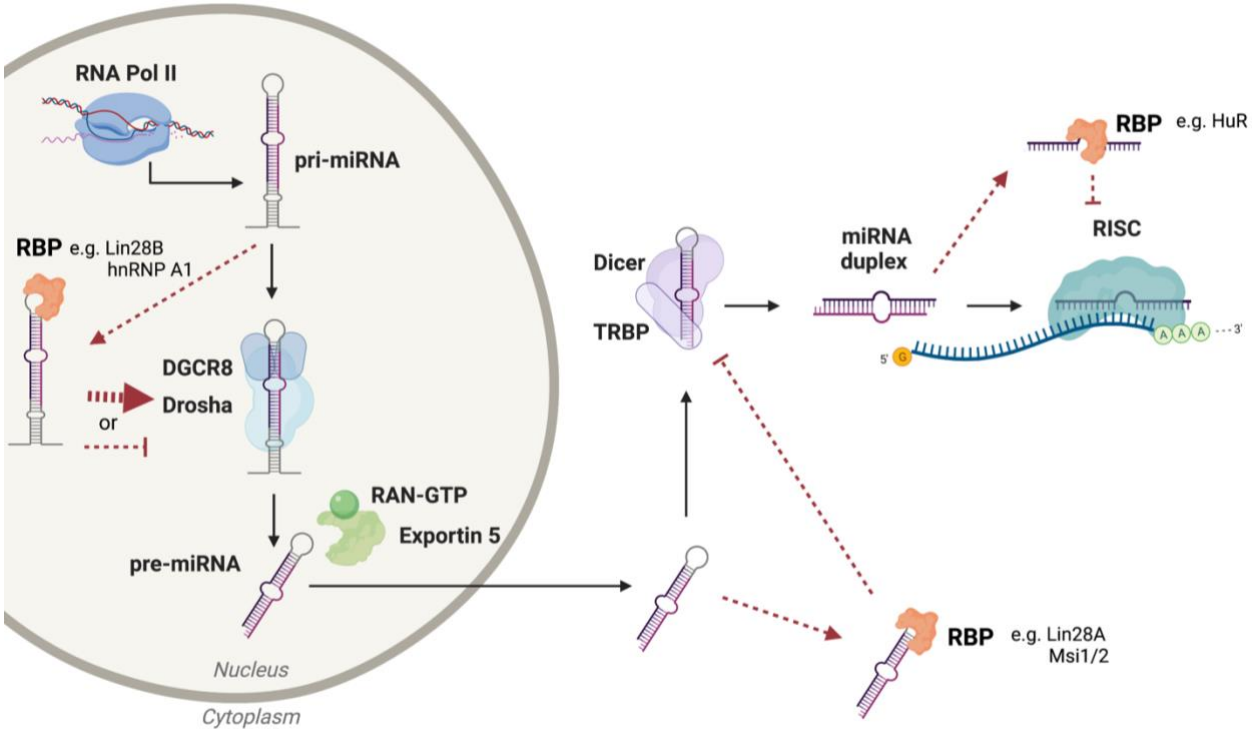


Figure 1.2. Role of RBPs in regulating miRNA biogenesis. RBP binding to pri- or pre-miRNA can enhance or inhibit further processing. Created with BioRender.com.

In 2017, the Meister group utilized a proteomics-based pulldown approach to identify RBPs that specifically interacted with pre-miRNA hairpins to discover potential regulatory pathways. From this large-scale study, ~180 RBPs were found to interact specifically with a unique subset of the 72 pre-miRNA hairpin baits used. Several interactions were validated using biochemical methods, including reciprocal pulldowns coupled to Western blot, Northern blot, and qPCR. The functional consequences of these interactions were explored by knocking down or overexpressing selected RBPs, which revealed RBP-dependent changes to mature miRNA levels. This study presents a large number of pre-miRNA-RBP interactions that require additional characterization.[28]

Let-7 was the second miRNA to be discovered in *C. elegans* and was characterized as a heterochronic gene required for normal larval development.[43-44] Garnering much attention due to the high degree of conservation of let-7 across phyla, it has become one of the best characterized miRNAs to date.[45] Throughout evolution, duplication events led to distinct members of the let-7 family in higher organisms.[46] In humans, there are 13 distinct let-7 sequences, one of which, let-7a is identical across the animal kingdom, from *C. elegans* to humans.[45, 47] Remarkably, the let-7 seed sequence, GAGGUAG, is identical across the entire family in humans as well as across species.[48]

As with let-7, the RBP Lin28 was first discovered as a player in developmental timing in *C. elegans* and later found to promote pluripotency across several species, including *Drosophila*, *Xenopus*, mouse, and humans, by acting as a regulator of the let-7 family of miRNA.[49-50] Higher mammals express two paralogs, Lin28A and Lin28B. The human paralogs, which share 65% sequence identity, containing nearly identical RNA-binding domains (RBDs), an N-terminal cold

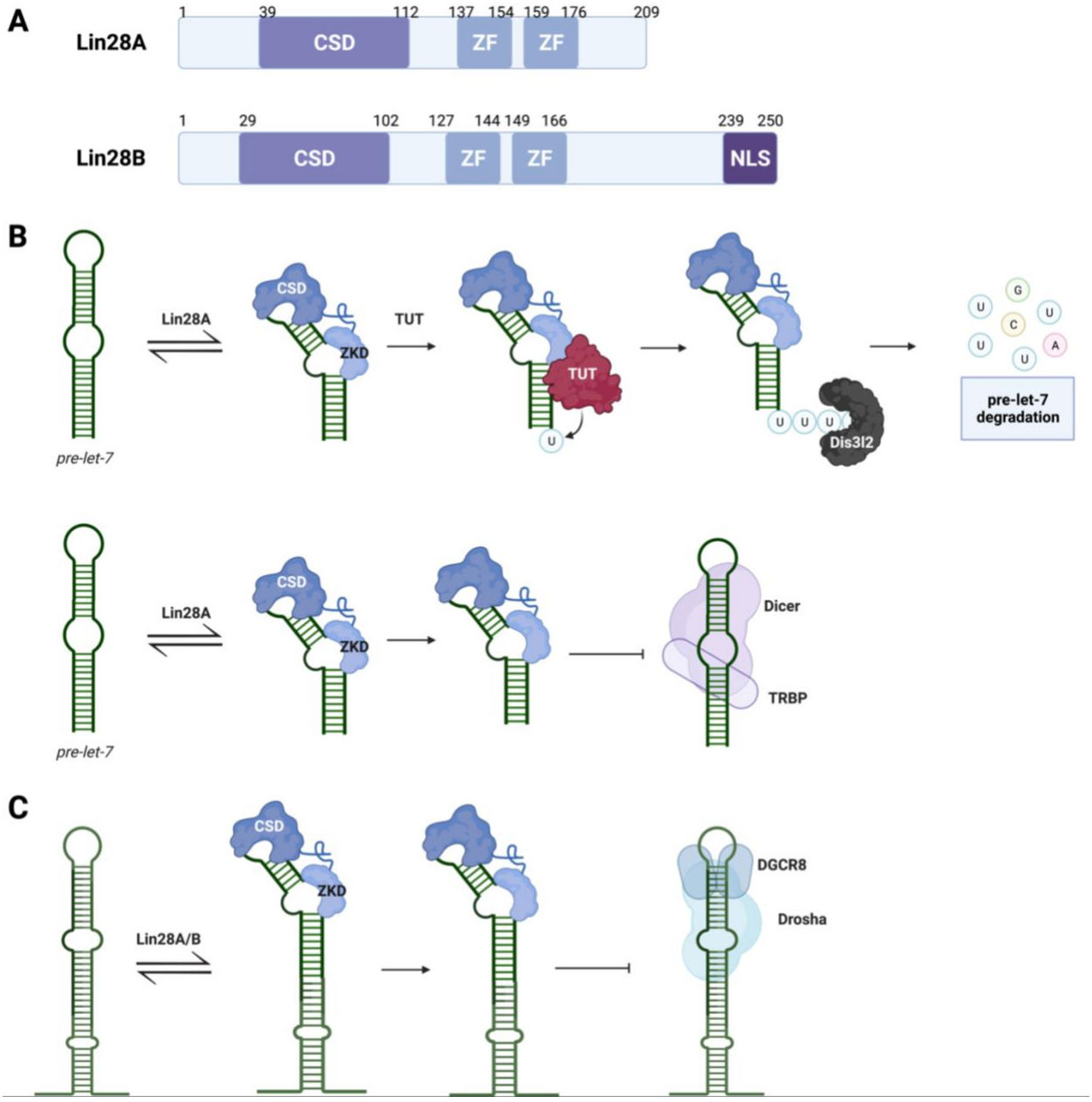


Figure 1.3 The let-7/Lin28 interaction. (A) Domain organization of Lin28A and Lin28B, highlighting the cold shock domain (CSD) and zinc knuckle domain (ZKD) which is composed of tandem zinc fingers (ZF). **(B)** Functional interaction between let-7 and Lin2A8A and **(C)** let-7 and Lin28B. Created with BioRender.com.

shock domain (CSD) and C-terminal zinc knuckle domain (ZKD), connected by a flexible linker (Fig. 1.3A).[51] While Lin28A is primarily localized to the cytoplasm despite possessing a putative bipartite nucleolar localization sequence (NoLS), Lin28B functions predominantly in the nucleus owing to its C-terminal nuclear localization signal (Fig. 1.3A).[52]

The Lin28 proteins bind to the terminal loop of either pri- or pre-let-7 through interactions between the RNA and both the CSD and ZKD. The CSD, which remodels the let-7 loop upon binding, binds with high affinity but low sequence specificity, showing a preference for a GNGAY motif found in some but not all the let-7 family members.[53-54] The ZKD, composed of tandem retroviral-type CCHC zinc finger motifs, selectively binds to the conserved let-7 GGAG motif, but with lower affinity than the CSD.[50, 55-56] This binding to the terminal loop of let-7 precursors inhibits processing by Drosha in the nucleus and Dicer in the cytoplasm (Fig. 1.3B and 1.3C).[57-60]

Uniquely, Lin28A functions in the cytoplasm to reduce let-7 levels by inducing 3' end oligouridylation by the terminal uridyltransferases (TUTases) TUT4/Zcchc11 and TUT7/Zcchc6.[58, 61-63] Through structural and biochemical studies, Wang *et al*., demonstrated that the ZKD of Lin28A is required for recruitment of TUT4 and TUT7.[64] The processive addition of a polyuridine tail to the 3' end of let-7 signals the RNA for degradation by the 3'-5' exonuclease, Dis3l2 (Fig. 1.3B).[65-66]

In mammals, let-7 is virtually undetectable in embryonic stem cells (ESCs) owing to high levels of Lin28 expression; and upon cellular differentiation during development, Lin28 expression is turned off causing a corresponding increase in let-7 levels.[67-69] Most differentiated tissues, with the exception of skeletal and cardiac muscle, exhibit no Lin28A expression and maintain high levels of let-7.[70] Due to their role in depressing the expression of key proteins, such as c-Myc, Ras, and HMGA2, involved in development, pluripotency maintenance, muscle formation, cell adhesion, and proliferation, members of the let-7 family of miRNA are classified as tumor suppressors.[14, 71-75] Aberrant expression of Lin28 in differentiated cells in adulthood is found in ~15% of tumors, including glioblastoma, ovarian, gastric, prostate, and breast cancers.[76-78] Accordingly, the let-7/Lin28 interaction was recognized as a promising target for the treatment

of cancers, inspiring numerous drug discovery efforts that will be discussed in greater detail in later chapters.

1.3.1 hnRNPA1

Heterogeneous nuclear protein A1 (hnRNP A1) is a nucleo-cytoplasmic shuttling protein involved in the regulation of gene expression across a variety of pathways. hnRNP A1 binds nucleic acids through tandem RNA recognition motif (RRM) domains composed of the highly conserved RNP-1 and RNP-2 motifs, collectively termed Unwinding Protein 1 (UP1), and possesses a C-terminal glycine-rich low complexity region, which is contains the M9 sequence responsible for nuclear export and import.[79] HnRNP A1 acts as transcriptional regulator through direct promoter interactions and association with transcription factors, plays a role in mRNA splicing, nuclear export, and turnover, as well as IRES-mediated translation and telomere maintenance (Fig. 1.4).[80] In addition, hnRNP A1 has been characterized as a regulator of miR-18a and let-7 biogenesis. This function of hnRNP A1 was serendipitously discovered by the Cáceres group as they endeavored to identify cellular RNA targets of hnRNP A1. By using the cross-linking and immunoprecipitation (CLIP) method to identify sequences bound by hnRNP A1, 200 sequences were identified, one of which was a miRNA. hnRNP A1 was found to specifically and exclusively facilitate processing of pri-miR-18a, despite miR-18a being located within the polycistronic miR-17-92 cluster.[7] Notably, the miR-17-92 cluster, which contains six distinct pre-miRNA hairpins, is involved in the regulation of proliferation, cell cycle, apoptosis, and differentiation and amplification of the miR-17 cluster has been linked to several cancers, including human B-cell lymphomas and colon cancers.[81-83] MiR-18a targets many genes involved in cancers, including IRF2, SOX6, PTEN, and CDC42. Intriguingly, miR-18a has been

characterized as an onco-miR, contributing to the progression of certain cancers, as well as a tumor suppressor in other tumor types by inhibiting proliferation and epithelial-mesenchymal transition (EMT). In mesothelioma and lung, cervical, prostate, and gastric cancers, miR-18a promoted disease progression, whereas in colorectal and breast cancers miR-18a was found to inhibit cell division, migration, and invasion.[83]

Using SELEX, the consensus binding sequence of hnRNP A1 was determined to be UAGGGA/U.[84] Further investigation revealed that hnRNP A1 binds in two regions of pri-miR-

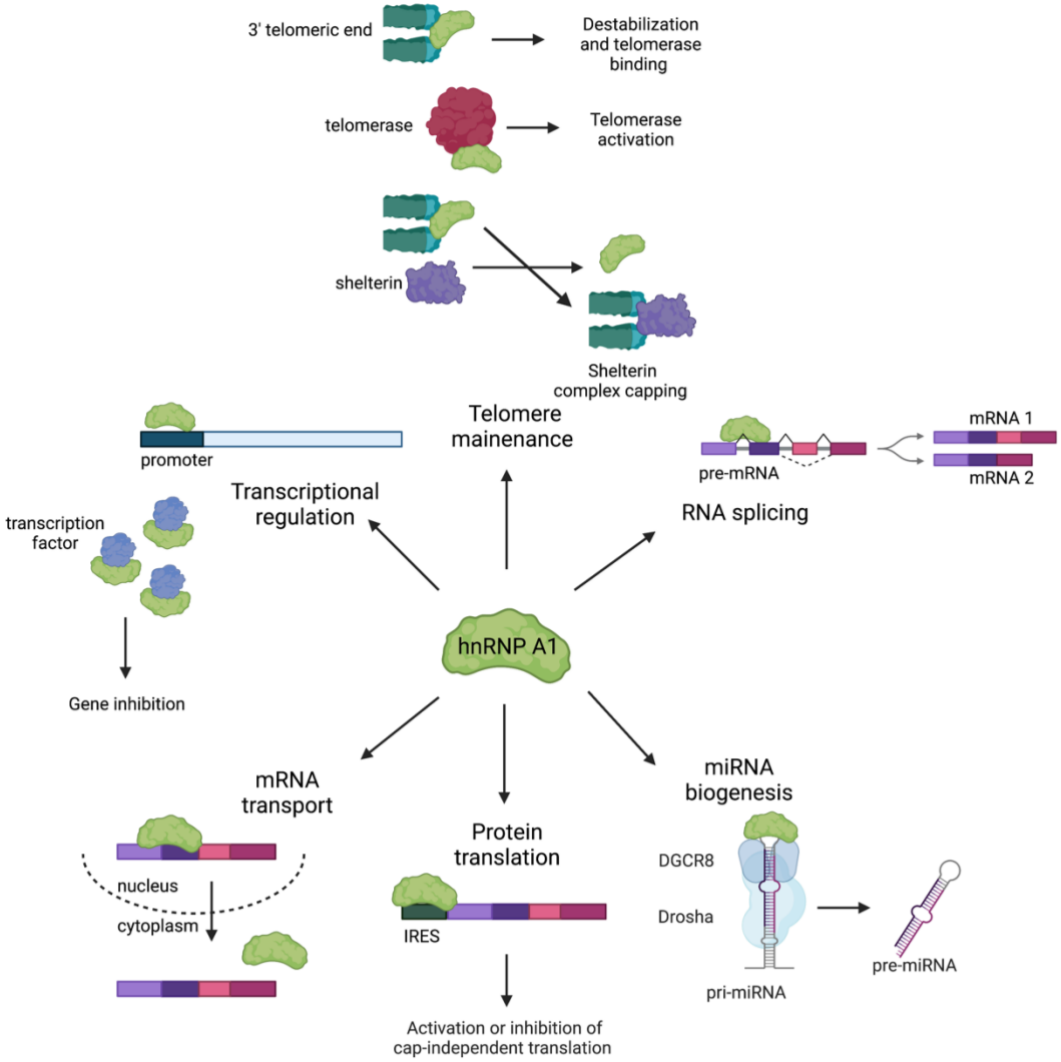


Figure 1.4 Functional roles of hnRNP A1. Depiction of the role of hnRNP A1 in telomerase maintenance and the regulation of transcription, RNA splicing, mRNA transport, translation, and miRNA biogenesis. Figure adapted from Clarke et al., *Front. Mol. Biosci.* 2021. Created with BioRender.com.

18a, to a UAG motif in the terminal loop as previously observed in the CLIP data, as well as a

UAG motif in an internal loop at the base of the stem that is observed upon hnRNP A1 binding.[7, 85] It was proposed that by binding to the terminal loop of miR-18a, hnRNP A1 induces a relaxation change in the stem region, generating an optimal conformation for Drosha cleavage.[85] Additionally, the Cáceres group utilized RNA chromatography coupled with mass spectrometry to identify auxiliary factors involved in the regulation of other miRNAs that, similar to miR-18a, contain conserved terminal loop regions and found hnRNP A1 specifically bound to other pri-miRNAs, including pre-let-7a-1 and pri-miR-101-1.[85] In line with hnRNP A1 binding preference, pre-let-7a-1 contains UAGGGU in the terminal loop.[85] Follow-up studies revealed that hnRNP A1 binding to the pri-let-7a-1 terminal loop inhibits processing to pre-let-7a-1 by Drosha.[9]

1.3.2 Musashi1/2

The Musashi (Msi) RBP was initially identified in *Drosophila* due to its role in asymmetric cell division, stem cell function, and cell fate determination via translational repression and activation.[86] Vertebrates express two evolutionarily conserved Msi orthologs, Msi1 and Msi2.[86] Common among all Msi family members are two tandem N-terminal RRMs, RRM1 and RRM2, which contain the highly conserved RNP-1 and RNP-2 motifs. In humans, Msi2 is 69% identical to Msi1 with their RRMs containing 87% identity and both bind to a similar UAG motif, resulting in some functional redundancy.[87-90]

Despite their structural similarity, different roles for Msi1 and Msi2 have been reported. In mammalian cells, Msi1 is highly expressed in neural stem/progenitor cells (NS/PCs) and other somatic stem cells and controls the stemness and cell fate by acting as a translational regulator. In some cases, Msi1 represses translation by binding to the 3' UTR of target mRNA, sequestering poly(A) binding protein (PABP), blocking PABP from interacting with eIF4G, which in turn inhibits the formation of the 80s ribosome complex.[86, 91] In particular, Msi1 is known to inhibit the

translation of transcripts such as *m-numb* and *CDKN1A* in mammals.[92-93] Conversely, Msi1 binding to some transcripts is stabilizing and enhances their translation.[94] Further exploration of Msi1 targets identified additional putative targets, 735 of which were up-regulated and 31 of which were down-regulated upon Msi1 knockdown in 5637 bladder carcinoma cells.[95]

Another reported role of Msi1, is a synergistic relationship with Lin28 in the inhibition of let-7 biogenesis. Kawahara *et al.* found that Msi1 interacts with Lin28, in the nucleus independent of RNA and in the cytoplasm dependent on binding to mRNA. Notably, Msi2 was not found to interact with Lin28. Msi1 enhances Lin28 localization to the nucleus, aiding in inhibiting let-7 biogenesis. Specifically, Lin28-mediated inhibition of pri-miR-98 cleavage by Drosha was enhanced by Msi1 in a dose-dependent manner.[91]

Unlike Msi1, Msi2 is expressed in a wider variety of cell types and tissues and is strongly associated with maintenance of stemness in neuronal and hematopoietic stem cells.[86] Msi2 has been studied as a driver of oncogenesis and pathogenic progression of myeloid leukemia.[86] Less is known about the targets and mechanisms of Msi2-mediated regulation, but several studies have recently sought to identify RNA targets of Msi2, specifically in the context of AML and related stem cells (leukemia and hematopoietic stem cells).[96-98] Interestingly, when mapping Msi2 mRNA targets, it was found that Msi2 RNA binding activity changes upon differentiation and is significantly increased in leukemia versus hematopoietic stem cells.[98]

Msi2 has also been linked to the regulation of miRNA biogenesis, particularly miR-7. To elucidate determinants of tissue-specific miRNA levels in brain tissues, Choudhury *et al.* discovered that, in nonneuronal cells, Msi2 is recruited to the terminal loop of pri-miRNA by another RBP, Hu antigen R (HuR). This interaction stabilizes the pri-miR-7 structure and inhibits subsequent processing to mature miR-7.[99] In addition to miR-7, miR-505, miR-92a-1, and miR-

224 were also sensitive to Msi2 and HuR knockdown, indicating there are regulatory mechanisms that require further exploration. More recently, Msi1 and Msi2 were found to strongly interact with pre-miR-18a and pre-miR-20a, although further studies are required to characterize the functional interaction.[28]

1.4 Conclusion

Given the importance of maintaining miRNA levels for proper cellular function and contribution to disease states, it is imperative to characterize pathways that regulate miRNA biogenesis and levels. Several RBPs, including Lin28A/B, hnRNP A1, and Msi1/2, have been studied to various extents for their role in modulating miRNA biogenesis. Due to their influence on the levels of disease-relevant miRNAs, and considering the challenges associated with targeting RNA itself, restoring miRNA levels through the targeting of miR-RBP interactions has emerged as a promising therapeutic strategy. Efforts not only to validate, but also target miRNA-RBP interactions has revealed a need for technologies that detect of such interactions for the validation putative binding partners and enable subsequent drug discovery.

1.5 References

1. Friedman, R. C.; Farh, K. K. H.; Burge, C. B.; Bartel, D. P., Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **2009,** *19* (1), 92-105.

2. Ha, M.; Kim, V. N., Regulation of microRNA biogenesis. *Nat Rev Mol Cell Bio* **2014,** *15* (8), 509-524.

3. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S., miRBase: from microRNA sequences to function. *Nucleic Acids Res* **2019,** *47* (D1), D155-D162.

4. Filipowicz, W.; Bhattacharyya, S. N.; Sonenberg, N., Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **2008,** *9* (2), 102-14.

5. Bartel, D. P., MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **2004,** *116* (2), 281-97.

6. Davis-Dusenbery, B. N.; Hata, A., Mechanisms of control of microRNA biogenesis. *J Biochem* **2010,** *148* (4), 381-92.

7. Guil, S.; Caceres, J. F., The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Structural & Molecular Biology* **2007,** *14* (7), 591-596.

8. Hentze, M. W.; Castello, A.; Schwarzl, T.; Preiss, T., A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* **2018,** *19* (5), 327-341.

9. Michlewski, G.; Caceres, J. F., Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nature Structural & Molecular Biology* **2010,** *17* (8), 1011-1128.

10. Min, K. W.; Jo, M. H.; Shin, S.; Davila, S.; Zealy, R. W.; Kang, S. I.; Lloyd, L. T.; Hohng, S.; Yoon, J. H., AUF1 facilitates microRNA-mediated gene silencing. *Nucleic Acids Res* **2017,** *45* (10), 6064-6073.

11. van Kouwenhove, M.; Kedde, M.; Agami, R., MicroRNA regulation by RNA-binding proteins and its implications for cancer. *Nature Reviews Cancer* **2011,** *11* (9), 644-656.

12. Lin, S. B.; Gregory, R. I., MicroRNA biogenesis pathways in cancer. *Nature Reviews Cancer* **2015,** *15* (6), 321-333.

13. Li, Y.; Kowdley, K. V., MicroRNAs in common human diseases. *Genomics Proteomics Bioinformatics* **2012,** *10* (5), 246-53.

14. Lee, Y. S.; Dutta, A., MicroRNAs in cancer. *Annu Rev Pathol* **2009,** *4*, 199-227.

15. Ardekani, A. M.; Naeini, M. M., The role of MicroRNAs in human diseases. *Avicenna Journal of Medical Biotechnology* **2010,** *2*, 161-179.

16. Lee, Y.; Kim, M.; Han, J.; Yeom, K. H.; Lee, S.; Baek, S. H.; Kim, V. N., MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **2004,** *23* (20), 4051-60.

17. Han, J.; Lee, Y.; Yeom, K. H.; Kim, Y. K.; Jin, H.; Kim, V. N., The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **2004,** *18* (24), 3016-27.

18. Zeng, Y.; Cullen, B. R., Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* **2004,** *32* (16), 4776-85.

19. Gulyaeva, L. F.; Kushlinskiy, N. E., Regulatory mechanisms of microRNA expression. *J Transl Med* **2016,** *14* (1), 143.

20. Yoda, M.; Kawamata, T.; Paroo, Z.; Ye, X.; Iwasaki, S.; Liu, Q.; Tomari, Y., ATP-dependent human RISC assembly pathways. *Nat Struct Mol Biol* **2010,** *17* (1), 17-23.

21. Chendrimada, T. P.; Gregory, R. I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R., TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **2005,** *436* (7051), 740-4.

22. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C., Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **2018,** *9*, 402.

23. Hock, J.; Meister, G., The Argonaute protein family. *Genome Biol* **2008,** *9* (2), 210.

24. Forman, J. J.; Legesse-Miller, A.; Coller, H. A., A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proc Natl Acad Sci U S A* **2008,** *105* (39), 14879-84.

25. Zhang, J.; Zhou, W.; Liu, Y.; Liu, T.; Li, C.; Wang, L., Oncogenic role of microRNA-532-5p in human colorectal cancer via targeting of the 5'UTR of RUNX3. *Oncol Lett* **2018,** *15* (5), 7215-7220.

26. Jonas, S.; Izaurralde, E., Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **2015,** *16* (7), 421-33.

27. Jo, M. H.; Shin, S.; Jung, S. R.; Kim, E.; Song, J. J.; Hohng, S., Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. *Mol Cell* **2015,** *59* (1), 117-24.

28. Treiber, T.; Treiber, N.; Plessmann, U.; Harlander, S.; Daiss, J. L.; Eichner, N.; Lehmann, G.; Schall, K.; Urlaub, H.; Meister, G., A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Molecular Cell* **2017,** *66* (2), 270-284.

29. He, L.; He, X.; Lim, L. P.; de Stanchina, E.; Xuan, Z.; Liang, Y.; Xue, W.; Zender, L.; Magnus, J.; Ridzon, D.; Jackson, A. L.; Linsley, P. S.; Chen, C.; Lowe, S. W.; Cleary, M. A.; Hannon, G. J., A microRNA component of the p53 tumour suppressor network. *Nature* **2007,** *447* (7148), 1130-4.

30. Raver-Shapira, N.; Marciano, E.; Meiri, E.; Spector, Y.; Rosenfeld, N.; Moskovits, N.; Bentwich, Z.; Oren, M., Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Mol Cell* **2007,** *26* (5), 731-43.

31. Tarasov, V.; Jung, P.; Verdoodt, B.; Lodygin, D.; Epanchintsev, A.; Menssen, A.; Meister, G.; Hermeking, H., Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell Cycle* **2007,** *6* (13), 1586-93.

32. Bracken, C. P.; Gregory, P. A.; Kolesnikoff, N.; Bert, A. G.; Wang, J.; Shannon, M. F.; Goodall, G. J., A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res* **2008,** *68* (19), 7846-54.

33. Pawlicki, J. M.; Steitz, J. A., Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **2008,** *182* (1), 61-76.

34. Kim, Y. K.; Heo, I.; Kim, V. N., Modifications of small RNAs and their associated proteins. *Cell* **2010,** *143* (5), 703-9.

35. Auyeung, V. C.; Ulitsky, I.; McGeary, S. E.; Bartel, D. P., Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. *Cell* **2013,** *152* (4), 844-58.

36. MacRae, I. J.; Zhou, K.; Doudna, J. A., Structural determinants of RNA recognition and cleavage by Dicer. *Nat Struct Mol Biol* **2007,** *14* (10), 934-40.

37. Luo, Q. J.; Zhang, J.; Li, P.; Wang, Q.; Zhang, Y.; Roy-Chaudhuri, B.; Xu, J.; Kay, M. A.; Zhang, Q. C., RNA structure probing reveals the structural basis of Dicer binding and cleavage. *Nat Commun* **2021,** *12* (1), 3397.

38. Vermeulen, A.; Behlen, L.; Reynolds, A.; Wolfson, A.; Marshall, W. S.; Karpilow, J.; Khvorova, A., The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA* **2005,** *11* (5), 674-82.

39. Ciafre, S. A.; Galardi, S., microRNAs and RNA-binding proteins A complex network of interactions and reciprocal regulations in cancer. *Rna Biology* **2013,** *10* (6), 935-943.

40. Choudhury, N. R.; Michlewski, G., Terminal loop-mediated control of microRNA biogenesis. *Biochem Soc Trans* **2012,** *40* (4), 789-93.

41. Connerty, P.; Ahadi, A.; Hutvagner, G., RNA Binding Proteins in the miRNA Pathway. *Int J Mol Sci* **2015,** *17* (1).

42. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A. C.; Munschauer, M.; Ulrich, A.; Wardle, G. S.; Dewell, S.; Zavolan, M.; Tuschl, T., Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010,** *141* (1), 129-41.

43. Reinhart, B. J.; Slack, F. J.; Basson, M.; Pasquinelli, A. E.; Bettinger, J. C.; Rougvie, A. E.; Horvitz, H. R.; Ruvkun, G., The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **2000,** *403* (6772), 901-6.

44. Slack, F. J.; Basson, M.; Liu, Z.; Ambros, V.; Horvitz, H. R.; Ruvkun, G., The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol Cell* **2000,** *5* (4), 659-69.

45. Boyerinas, B.; Park, S. M.; Hau, A.; Murmann, A. E.; Peter, M. E., The role of let-7 in cell differentiation and cancer. *Endocr Relat Cancer* **2010,** *17* (1), F19-36.

46. Hertel, J.; Bartschat, S.; Wintsche, A.; Otto, C.; Students of the Bioinformatics Computer, L.; Stadler, P. F., Evolution of the let-7 microRNA family. *RNA Biol* **2012,** *9* (3), 231-41.

47. Ruby, J. G.; Jan, C.; Player, C.; Axtell, M. J.; Lee, W.; Nusbaum, C.; Ge, H.; Bartel, D. P., Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell* **2006,** *127* (6), 1193-207.

48. Lee, H.; Han, S.; Kwon, C. S.; Lee, D., Biogenesis and regulation of the let-7 miRNAs and their functional implications. *Protein Cell* **2016,** *7* (2), 100-13.

49. Ambros, V.; Horvitz, H. R., Heterochronic mutants of the nematode Caenorhabditis elegans. *Science* **1984,** *226* (4673), 409-16.

50. Moss, E. G.; Tang, L., Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev Biol* **2003,** *258* (2), 432-42.

51. Mayr, F.; Heinemann, U., Mechanisms of Lin28-mediated miRNA and mRNA regulation--a structural and functional perspective. *Int J Mol Sci* **2013,** *14* (8), 16532-53.

52. Piskounova, E.; Polytarchou, C.; Thornton, J. E.; LaPierre, R. J.; Pothoulakis, C.; Hagan, J. P.; Iliopoulos, D.; Gregory, R. I., Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* **2011,** *147* (5), 1066-79.

53. Mayr, F.; Schutz, A.; Doge, N.; Heinemann, U., The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Res* **2012,** *40* (15), 7492-506.

54. Ustianenko, D.; Chiu, H.-S.; Treiber, T.; Weyn-Vanhentenryck, S. M.; Treiber, N.; Meister, G.; Sumazin, P.; Zhang, C., LIN28 selectively modulates a subclass of let-7 microRNAs. *Mol. Cell* **2018,** *71*, 271-283.

55. Nam, Y.; Chen, C.; Gregory, R. I.; Chou, J. J.; Sliz, P., Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **2011,** *147* (5), 1080-91.

56. Loughlin, F. E.; Gebert, L. F.; Towbin, H.; Brunschweiger, A.; Hall, J.; Allain, F. H., Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat Struct Mol Biol* **2011,** *19* (1), 84-9.

57. Rybak, A.; Fuchs, H.; Smirnova, L.; Brandt, C.; Pohl, E. E.; Nitsch, R.; Wulczyn, F. G., A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol* **2008,** *10* (8), 987-93.

58. Heo, I.; Joo, C.; Cho, J.; Ha, M.; Han, J. J.; Kim, V. N., Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell* **2008,** *32* (2), 276-284.

59. Viswanathan, S. R.; Daley, G. Q.; Gregory, R. I., Selective blockade of microRNA processing by Lin28. *Science* **2008,** *320* (5872), 97-100.

60. Newman, M. A.; Thomson, J. M.; Hammond, S. M., Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **2008,** *14* (8), 1539-49.

61. Hagan, J. P.; Piskounova, E.; Gregory, R. I., Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol* **2009,** *16* (10), 1021-5.

62. Heo, I.; Joo, C.; Kim, Y. K.; Ha, M.; Yoon, M. J.; Cho, J.; Yeom, K. H.; Han, J.; Kim, V. N., TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell* **2009,** *138* (4), 696-708.

63. Thornton, J. E.; Chang, H. M.; Piskounova, E.; Gregory, R. I., Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *Rna* **2012,** *18* (10), 1875-1885.

64. Wang, L.; Nam, Y.; Lee, A. K.; Yu, C.; Roth, K.; Chen, C.; Ransey, E. M.; Sliz, P., LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. *Cell Reports* **2017,** *18* (11), 2664-2675.

65. Chang, H. M.; Triboulet, R.; Thornton, J. E.; Gregory, R. I., A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature* **2013,** *497* (7448), 244-+.

66. Ustianenko, D.; Hrossova, D.; Potesil, D.; Chalupnikova, K.; Hrazdilova, K.; Pachernik, J.; Cetkovska, K.; Uldrijan, S.; Zdrahal, Z.; Vanacova, S., Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **2013,** *19* (12), 1632-8.

67. Schulman, B. R.; Esquela-Kerscher, A.; Slack, F. J., Reciprocal expression of lin-41 and the microRNAs let-7 and mir-125 during mouse embryogenesis. *Dev Dyn* **2005,** *234* (4), 1046-54.

68. Thomson, J. M.; Parker, J.; Perou, C. M.; Hammond, S. M., A custom microarray platform for analysis of microRNA gene expression. *Nat Methods* **2004,** *1* (1), 47-53.

69. Thomson, J. M.; Newman, M.; Parker, J. S.; Morin-Kensicki, E. M.; Wright, T.; Hammond, S. M., Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev* **2006,** *20* (16), 2202-7.

70. Polesskaya, A.; Cuvellier, S.; Naguibneva, I.; Duquet, A.; Moss, E. G.; Harel-Bellan, A., Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. *Gene Dev* **2007,** *21* (9), 1125-1138.

71. Wang, X. R.; Cao, L.; Wang, Y. Y.; Wang, X. F.; Liu, N.; You, Y. P., Regulation of let-7 and its target oncogenes (Review). *Oncol Lett* **2012,** *3* (5), 955-960.

72. Johnson, S. M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K. L.; Brown, D.; Slack, F. J., RAS is regulated by the let-7 MicroRNA family. *Cell* **2005,** *120* (5), 635-647.

73. Mayr, C.; Hemann, M. T.; Bartel, D. P., Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **2007,** *315* (5818), 1576-9.

74. Sampson, V. B.; Rong, N. H.; Han, J.; Yang, Q.; Aris, V.; Soteropoulos, P.; Petrelli, N. J.; Dunn, S. P.; Krueger, L. J., MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Res* **2007,** *67* (20), 9762-70.

75. Worringer, K. A.; Rand, T. A.; Hayashi, Y.; Sami, S.; Takahashi, K.; Tanabe, K.; Narita, M.; Srivastava, D.; Yamanaka, S., The let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. *Cell Stem Cell* **2014,** *14* (1), 40-52.

76. Shyh-Chang, N.; Daley, G. Q., Lin28: Primal Regulator of Growth and Metabolism in Stem Cells. *Cell Stem Cell* **2013,** *12* (4), 395-406.

77. Viswanathan, S. R.; Powers, J. T.; Einhorn, W.; Hoshida, Y.; Ng, T. L.; Toffanin, S.; O'Sullivan, M.; Lu, J.; Phillips, L. A.; Lockhart, V. L.; Shah, S. P.; Tanwar, P. S.; Mermel, C. H.; Beroukhim, R.; Azam, M.; Teixeira, J.; Meyerson, M.; Hughes, T. P.; Llovet, J. M.; Radich, J.; Mullighan, C. G.; Golub, T. R.; Sorensen, P. H.; Daley, G. Q., Lin28 promotes transformation and is associated with advanced human malignancies. *Nat Genet* **2009,** *41* (7), 843-8.

78. Balzeau, J.; Menezes, M. R.; Cao, S. Y.; Hagan, J. P., The LIN28/let-7 Pathway in Cancer. *Front Genet* **2017,** *8*, 1-16.

79. Kooshapur, H.; Choudhury, N. R.; Simon, B.; Muhlbauer, M.; Jussupow, A.; Fernandez, N.; Jones, A. N.; Dallmann, A.; Gabel, F.; Camilloni, C.; Michlewski, G.; Caceres, J. F.; Sattler, M., Structural basis for terminal loop recognition and stimulation of pri-miRNA-18a processing by hnRNP A1. *Nat Commun* **2018,** *9* (1), 2479.

80. Jean-Philippe, J.; Paz, S.; Caputi, M., hnRNP A1: the Swiss army knife of gene expression. *Int J Mol Sci* **2013,** *14* (9), 18999-9024.

81. Ota, A.; Tagawa, H.; Karnan, S.; Tsuzuki, S.; Karpas, A.; Kira, S.; Yoshida, Y.; Seto, M., Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. *Cancer Res* **2004,** *64* (9), 3087-95.

82. Dews, M.; Homayouni, A.; Yu, D.; Murphy, D.; Sevignani, C.; Wentzel, E.; Furth, E. E.; Lee, W. M.; Enders, G. H.; Mendell, J. T.; Thomas-Tikhonenko, A., Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nat Genet* **2006,** *38* (9), 1060-5.

83. Shen, K.; Cao, Z.; Zhu, R.; You, L.; Zhang, T., The dual functional role of MicroRNA-18a (miR-18a) in cancer development. *Clin Transl Med* **2019,** *8* (1), 32.

84. Burd, C. G.; Dreyfuss, G., RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J* **1994,** *13* (5), 1197-204.

85. Michlewski, G.; Guil, S.; Semple, C. A.; Caceres, J. F., Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol Cell* **2008,** *32* (3), 383-93.

86. Sutherland, J. M.; McLaughlin, E. A.; Hime, G. R.; Siddall, N. A., The Musashi family of RNA binding proteins: master regulators of multiple stem cell populations. *Adv Exp Med Biol* **2013,** *786*, 233-45.

87. Lan, L.; Xing, M.; Kashipathy, M.; Douglas, J.; Gao, P.; Battaile, K.; Hanzlik, R.; Lovell, S.; Xu, L., Crystal and solution structures of human oncoprotein Musashi-2 N-terminal RNA recognition motif 1. *Proteins* **2020,** *88* (4), 573-583.

88. Ohyama, T.; Nagata, T.; Tsuda, K.; Kobayashi, N.; Imai, T.; Okano, H.; Yamazaki, T.; Katahira, M., Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. *Nucleic Acids Res* **2012,** *40* (7), 3218-31.

89. Sundar, J.; Matalkah, F.; Jeong, B.; Stoilov, P.; Ramamurthy, V., The Musashi proteins MSI1 and MSI2 are required for photoreceptor morphogenesis and vision in mice. *J Biol Chem* **2021,** *296*, 100048.

90. Zearfoss, N. R.; Deveau, L. M.; Clingman, C. C.; Schmidt, E.; Johnson, E. S.; Massi, F.; Ryder, S. P., A conserved three-nucleotide core motif defines Musashi RNA binding specificity. *J Biol Chem* **2014,** *289* (51), 35530-41.

91. Kawahara, H.; Okada, Y.; Imai, T.; Iwanami, A.; Mischel, P. S.; Okano, H., Musashi1 cooperates in abnormal cell lineage protein 28 (Lin28)-mediated let-7 family microRNA biogenesis in early neural differentiation. *J Biol Chem* **2011,** *286* (18), 16121-30.

92. Imai, T.; Tokunaga, A.; Yoshida, T.; Hashimoto, M.; Mikoshiba, K.; Weinmaster, G.; Nakafuku, M.; Okano, H., The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA. *Mol Cell Biol* **2001,** *21* (12), 3888-900.

93. Battelli, C.; Nikopoulos, G. N.; Mitchell, J. G.; Verdi, J. M., The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. *Mol Cell Neurosci* **2006,** *31* (1), 85-96.

94. de Sousa Abreu, R.; Sanchez-Diaz, P. C.; Vogel, C.; Burns, S. C.; Ko, D.; Burton, T. L.; Vo, D. T.; Chennasamudaram, S.; Le, S. Y.; Shapiro, B. A.; Penalva, L. O., Genomic analyses of musashi1 downstream targets show a strong association with cancer-related processes. *J Biol Chem* **2009,** *284* (18), 12125-35.

95. Nikpour, P.; Baygi, M. E.; Steinhoff, C.; Hader, C.; Luca, A. C.; Mowla, S. J.; Schulz, W. A., The RNA binding protein Musashi1 regulates apoptosis, gene expression and stress granule formation in urothelial carcinoma cells. *J Cell Mol Med* **2011,** *15* (5), 1210-24.

96. Park, S. M.; Deering, R. P.; Lu, Y.; Tivnan, P.; Lianoglou, S.; Al-Shahrour, F.; Ebert, B. L.; Hacohen, N.; Leslie, C.; Daley, G. Q.; Lengner, C. J.; Kharas, M. G., Musashi-2 controls cell fate, lineage bias, and TGF-beta signaling in HSCs. *J Exp Med* **2014,** *211* (1), 71-87.

97. Rentas, S.; Holzapfel, N.; Belew, M. S.; Pratt, G.; Voisin, V.; Wilhelm, B. T.; Bader, G. D.; Yeo, G. W.; Hope, K. J., Musashi-2 attenuates AHR signalling to expand human haematopoietic stem cells. *Nature* **2016,** *532* (7600), 508-511.

98. Nguyen, D. T. T.; Lu, Y.; Chu, K. L.; Yang, X.; Park, S. M.; Choo, Z. N.; Chin, C. R.; Prieto, C.; Schurer, A.; Barin, E.; Savino, A. M.; Gourkanti, S.; Patel, P.; Vu, L. P.; Leslie, C. S.;

Kharas, M. G., HyperTRIBE uncovers increased MUSASHI-2 RNA binding activity and differential regulation in leukemic stem cells. *Nat Commun* **2020,** *11* (1), 2026.

99. Choudhury, N. R.; de Lima Alves, F.; de Andres-Aguayo, L.; Graf, T.; Caceres, J. F.; Rappsilber, J.; Michlewski, G., Tissue-specific control of brain-enriched miR-7 biogenesis. *Genes Dev* **2013,** *27* (1), 24-38.

CHAPTER 2

Existing Methods for Detecting, Validating, and Manipulating RNA-Protein Interactions

Given mounting interest in mapping RPIs and, of late, inhibiting RPIs, the last two decades have seen an increase in the development of methodologies for detecting, validating, and manipulating RPIs both *in vitro* and *in cellulo*, many of which have been applied to the detection to study and manipulate pre-miRNA-protein interactions. *In vitro* methods, such as electrophoretic mobility shift assay (EMSA), surface plasmon resonance (SPR), fluorescence polarization (FP), and immunoprecipitation (IP), have long served as useful tools for measuring affinity and sequence specificity of RPIs of interest. In addition to FP, the cat-ELCCA technique developed in the Garner laboratory serve as important platforms for the detection of RPI manipulation by small molecules. More recently, cellular strategies, including fluorescence resonance energy transfer (FRET)-, protein complementation-, and immunoprecipitation (IP)-based methods, have emerged as powerful techniques. This chapter describes several important methods that enable detection of the interaction and inhibition of miRNA-RBP pairs, highlighting their advantages and limitations, and presents the ways in which the work detailed in this dissertation circumvents many of the challenges discussed herein.

2.1 *In Vitro* Methods Used in RNA-targeted Drug Discovery

As mentioned in Chapter 1, pre-miRNA-RBP interactions have gained traction as the targets of drug discovery efforts. Fluorescence polarization (FP)- and fluorescence resonance

energy transfer (FRET)-based assays have been optimized and utilized to enable the discovery of pre-miRNA-protein interaction inhibitors. In FP, binding and inhibition of an RPI is detectable due to changes in polarization of a fluorophore-labeled biomolecule rotating at differing speeds based on the size of the complex formed. For example, FP assays have been developed to detect binding of Lin28 to a truncated pre-let-7g or pre-let-7f-1 hairpin conjugated to a fluorophore to measure inhibition by small molecules.[1-3] However, FP assays are subject to high false hit rates due to intrinsic fluorescence of compounds, induction of static or dynamic fluorophore quenching, or light scattering produced by compound precipitation.[1]

FRET-based methods require conjugation of the RBP- and RNA-of-interest to complementary donor and acceptor fluorophores. This strategy was similarly applied to the identification of pre-let-7/Lin28 inhibitors by Roos *et al.* and Lim *et al*. Roos *et al.* labeled Lin28 with EGFP and a FRET acceptor on a truncated pre-let-7a-2 hairpin, whereas Lim *et al.* directly modified Lin28 with a fluorophore via an unnatural amino acid in the flexible linker between the two Lin28 domains and labeled a truncated pre-let-7a-1 hairpin with a quencher molecule.[4-5] While capable of detecting direct RPIs and adaptable to HTS, these techniques did not utilize the full length pre-miRNA and are prone to interference by autofluorescent or naturally quenching molecules.

2.1.1 Methods Developed in the Garner Laboratory

Seeking to design an assay that eliminates some of the barriers faced when using fluorescence-based techniques, the Garner laboratory developed a high throughput, plate-based assay technology, catalytic enzyme-linked click chemistry assay (cat-ELCCA) for the screening of inhibitors of several biomolecular processes, including ghrelin *O*-acyltransferase (GOAT)[6] and Dicer activity[7-8], Lin28 binding to pre-let-7[9], and eIF4E-protein interactions[10]. For pre-let-7/Lin28

cat-ELCCA, biotinylated Lin28 is immobilized on a streptavidin-coated 384-well plate. Detection of the interaction between Lin28 and the RNA is detected via chemiluminescence signal generated by methyltetrazine (mTet)-labeled horseradish peroxidase (HRP), which is covalently conjugated to the RPI complex by the complementary click chemistry handle, 5'-*trans*-cyclooctene (TCO), in the pre-let-7 probe (Fig. 2.1). When competed with a small molecule inhibitor, the RNA probe is washed out of the well, resulting in a loss of signal.[9] The key advantages of cat-ELCCA are related to its enhanced sensitivity due to catalytic signal amplification, favorable assay statistics, and minimal compound interference compared to fluorescence-based assays.[11] While immobilizing the protein allows for stringent washing, resulting in significantly reduced background signal, this increases the cost and duration of the assay.

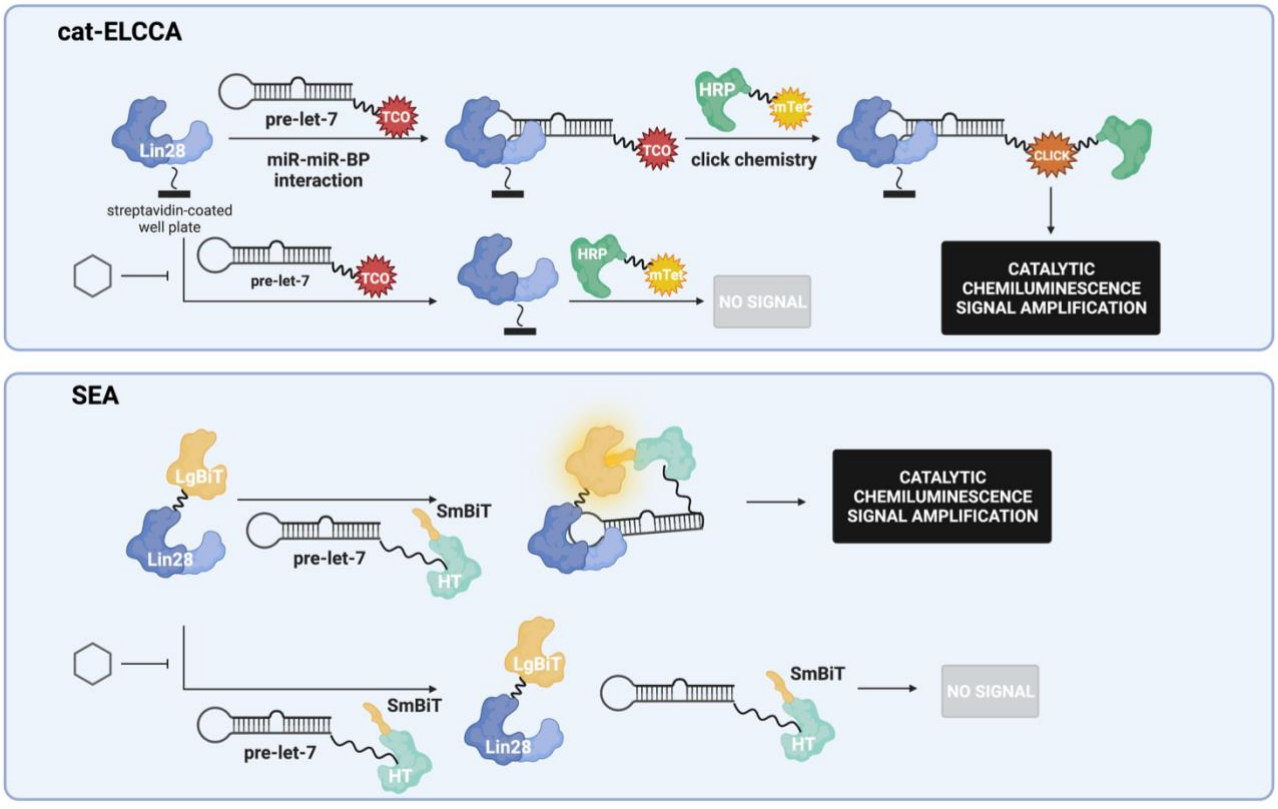


Figure 2.1. ***In vitro*** **RPI detection assays developed by the Garner laboratory.** Schematics of cat-ELCCA (top panel) and SEA (bottom panel) for the detection of the pre-let-7/Lin28 RPI.

Interested in eliminating the costly streptavidin-coated plates and washing steps required for cat-ELCCA, the Garner laboratory invented and optimized a novel homogenous click chemistry-based assay.[12] The resulting assay maintained the high sensitivity and HTS compatibility of cat-ELCCA, but was simplified into a one-pot system. This was achieved by utilizing protein complementation of a split luciferase was designed to enable generation of catalytic signal upon interaction between an RNA- and RBP-of-interest. The split NanoLuc enzyme, or NanoLuc Binary Technology (NanoBiT), is composed of a 18 kD large subunit (LgBiT) and an 11-amino peptide (SmBiT).[13] The BiTs were engineered to have low affinity for each other (K_d = 190 μM), allowing interaction between biomolecules fused to the individual BiTs to drive reassembly of the functional enzyme.[13] Additionally, the Lg-BiT/SmBiT interaction is reversible, allowing for the detection of interaction inhibition by small molecules.[13] In this assay, named Split Enzyme Assay (SEA), the RBP-of-interest fused to LgBiT, SmBiT fused to HaloTag, and a pre-miRNA probe containing a HaloTag ligand are mixed in a single tube. HaloTag, an engineered dehalogenase that covalently binds to chloroalkane-containing substrates[14], enables SmBiT labeling of the pre-miRNA probe and subsequent pre-miRNA-RBP interactions drive the reassembly of NanoLuc, allowing for the generation of catalytic signal amplification (Fig. 2.1).[12] Using pre-let-7/Lin28, this assay was shown to successfully detect interactions in 384-well plates and be adaptable to HTS with robust assay statistics.[12] While the advantages of SEA address limitations of cat-ELCCA and other fluorescence-based assay, it still suffers from the general limitations of *in vitro* methods.

2.1.2 General Limitations of In Vitro Methods

Despite variable approaches and strategies for detection, *in vitro* methods, with the exception of some immunoprecipitation-based techniques, commonly require a purified

biomolecule(s) (the RNA, RBP, or possibly both) and a non-native buffer medium. Purification of RNA and RBPs can be expensive, time consuming, and technically challenging. Since bacterial protein expression is widely used to generate large quantities of purified protein, the resultant protein will not contain post-translational modifications (PTMs) that can dictate RNA-binding activity. Similarly, *in vitro* transcribed RNAs will not be modified as they might be in a cellular context. Furthermore, purified systems could lack potentially required co-factors. Immunoprecipitation experiments which utilize affinity purification using an antibody for a protein does circumvent many of these limitations; however, even in lysate, it has been shown that non-physiological RNA-protein complexes can form.[15] In general, the use of non-physiological conditions, such as lysis and binding buffers, hinders the ability of *in vitro* techniques to accurately recapitulate the interaction as it occurs in the cell. Consequently, cellular methods for detecting RPIs serve an appealing alternative.

2.2 Cell-based Assays Used in RNA-targeted Drug Discovery

There are several key benefits of utilizing a cell-based platform for detecting RPIs and screening for inhibitors. Cellular assays have the potential to mirror physiological RPIs more closely and contain built-in selection for cellular permeability and activity, as well as cytotoxicity. There are many useful RPI detection techniques that are not covered in the present work, namely CLIP-based methods and incPRINT, as they are designed to primarily discover interactions or widely profile the binding preferences of an RBP- or RNA-of-interest and are not designed for HTS of small molecule inhibitors. Other methods of screening for miRNA inhibitors, in particular the dual luciferase assay, miRGlo, and microarrays, will also not be covered due to their focus on targeting the RNA species alone, either phenotypically or directly.

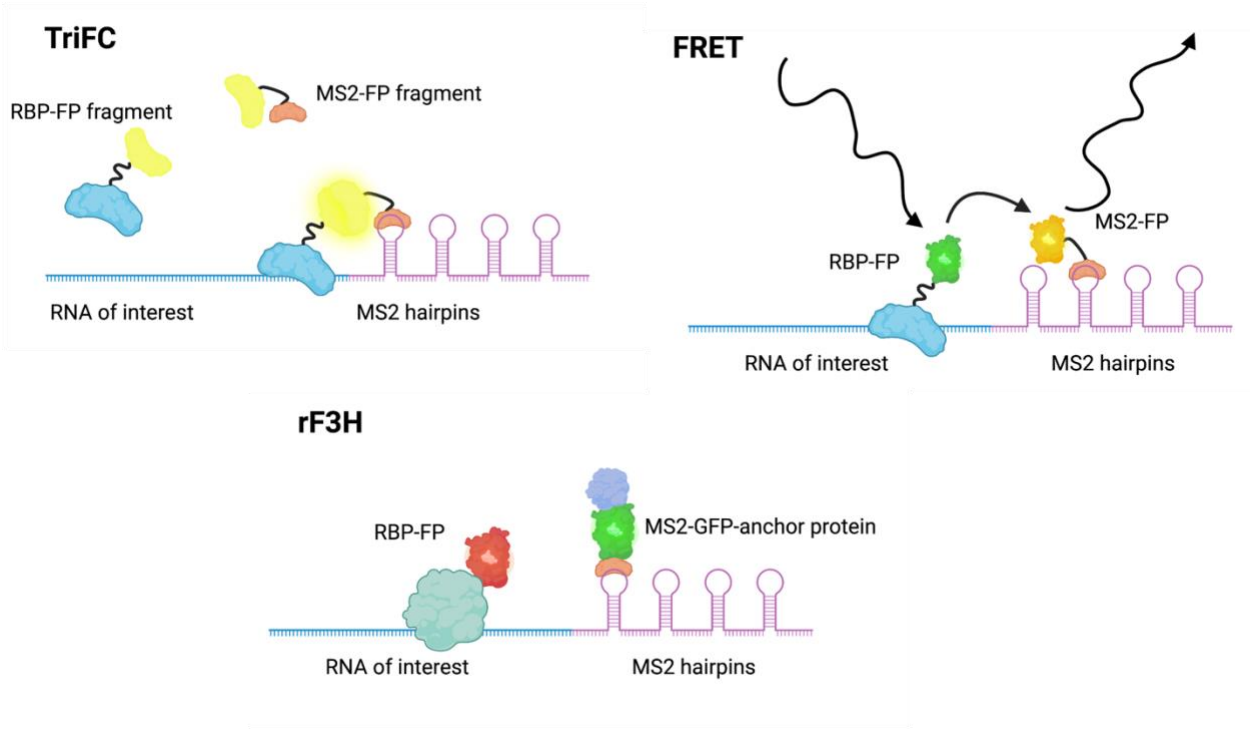


Figure 2.2. Cell-based RPI detection assays. Schematics of trimolecular fluorescence complementation (TriFC), fluorescence resonance energy transfer (FRET), and RNA fluorescence 3-hybrid (rF3H) assays.

2.2.1 Trimolecular Fluorescence Complementation (TriFC)

One strategy for visualizing RPIs in live cells is the use of fluorescence complementation. This approach, originally applied to the visualization of protein-protein interactions (PPIs) in cellular environments, is based on the ability of an interaction to facilitate the complementation of two fragments of a fluorescent protein. While labeling of the RBP with a protein fragment is easy and simply requires transgenic expression of a fusion protein, it is more challenging to label the RNA with the complementary protein fragment. To do so, the specific interaction between the RNA bacteriophage MS2 coat protein (MS2CP) and a specific stem-loop structure, the MS2 hairpin, is leveraged.[16-17] In TriFC for RPIs, MS2 hairpins are appended to the RNA-of-interest, which enables labeling of the RNA with the protein fragment in the cell via its fusion to the MS2CP. Formation of the tertiary complex between the MS2-labeled RNA and split fluorescent

protein-labeled MS2CP and RBP renders fluorescence signal (Fig. 2.2). Using this technique enables not only detection of interactions, but their localization in near native cellular conditions. According to a 2019 review, there have been no reported uses of the TriFC assay for screening, which is presumably due to low temperature requirements for fluorescence complementation, high background due to intrinsic affinity of the protein fragments, and the irreversible nature of complementation.[18]

2.2.2 FRET-based

As previously described, FRET requires the labeling of the RNA and the RBP with a molecule that can donate and one that can accept the energy transfer or quench fluorescence. One technique utilized a pan-RNA label SytoxOrange, which enabled visualization of all RNA interactions with a YFP-labeled RBP.[19] In order to visualize a specific RNA/RBP pair, however, the RNA-of-interest must be specifically labeled with a FRET component. To do so, Huranová *et al.* modified the RNA-of-interest with the MS2 stem-loop motif to enable recruitment of an MS2CP-labeled fluorescent protein. The compatible fluorescent protein pair when brought into close enough proximity via interaction of the RBP-CFP and the MS2CP-YFP labeled RNA-of-interest generates FRET signal (Fig. 2.2).[20] While FRET-based assays have been used to screen for small molecule inhibitors *in vitro*, cellular screening for RPI inhibitors using FRET has not been attempted.

2.2.3 rF3H

Recently, an RNA fluorescence three-hybrid (rF3H) method was developed to detect in cell RPIs. In rF3H, an MS2 stem-loop-labeled RNA-of-interest is anchored at a specific locus by a fusion protein composed of MS2CP, GFP, and an anchoring protein. Specific anchor proteins

can direct the locus of the trap, such as LacI binding to an integrated lac operon, Lamin B1 localizing to the nuclear lamina, or Coilin protein localizing to Cajal bodies. Subsequent interaction with a fluorescent protein fused RBP-of-interest can be detected by colocalization of the fluorescent signal (Fig. 2.2).[21] In this method, information about the native localization of the interaction may be lost depending on the anchor site chosen. However, enrichment of a signal at a specific locus could be advantageous for signal detection or for certain disease contexts.

2.3 Conclusion

The limitations of *in vitro* techniques used to screen for small molecule pre-miRNA-RBP inhibitors has highlighted the need for a cell-based platform for screening. The cellular methods discussed in this chapter rely on fluorescence outputs, which suffer from compound interference and require specialized microscopes for visualization. Furthermore, the use of the MS2 stem-loop significantly limits the utility of this method for the detection of highly processed RNAs. In the case of pre-miRNA, the MS2 stem-loop would be removed during or could disrupt miRNA biogenesis. Together, this constitutes a strong case for the development of a novel cell-based assay for the detection of direct interactions between RBPs and pre-miRNA hairpins.

2.4 References

1. Lightfoot, H. L.; Miska, E. A.; Balasubramanian, S., Identification of small molecule inhibitors of the Lin28-mediated blockage of pre-let-7g processing. *Org Biomol Chem* **2016,** *14* (43), 10208-10216.

2. Wang, L.; Rowe, R. G.; Jaimes, A.; Yu, C.; Nam, Y.; Pearson, D. S.; Zhang, J.; Xie, X.; Marion, W.; Heffron, G. J.; Daley, G. Q.; Sliz, P., Small-Molecule Inhibitors Disrupt let-7

Oligouridylation and Release the Selective Blockade of let-7 Processing by LIN28. *Cell Rep* **2018,** *23* (10), 3091-3101.

3. Borgelt, L.; Li, F.; Hommen, P.; Lampe, P.; Hwang, J.; Goebel, G. L.; Sievers, S.; Wu, P., Trisubstituted Pyrrolinones as Small-Molecule Inhibitors Disrupting the Protein-RNA Interaction of LIN28 and Let-7. *Acs Med Chem Lett* **2021,** *12* (6), 893-898.

4. Roos, M.; Pradere, U.; Ngondo, R. P.; Behera, A.; Allegrini, S.; Civenni, G.; Zagalak, J. A.; Marchand, J. R.; Menzi, M.; Towbin, H.; Scheuermann, J.; Neri, D.; Caflisch, A.; Catapano, C. V.; Ciaudo, C.; Hall, J., A Small-Molecule Inhibitor of Lin28. *ACS Chem Biol* **2016,** *11* (10), 2773-2781.

5. Lim, D.; Byun, W. G.; Koo, J. Y.; Park, H.; Park, S. B., Discovery of a Small-Molecule Inhibitor of Protein-MicroRNA Interaction Using Binding Assay with a Site-Specifically Labeled Lin28. *Journal of the American Chemical Society* **2016,** *138* (41), 13630-13638.

6. Garner, A. L.; Janda, K. D., cat-ELCCA: a robust method to monitor the fatty acid acyltransferase activity of ghrelin O-acyltransferase (GOAT). *Angew. Chem., Int. Ed.* **2010,** *49*, 9630-9634.

7. Lorenz, D. A.; Garner, A. L., A click chemistry-based microRNA maturation assay optimized for high-throughput screening. *Chem. Commun.* **2016,** (52), 8267-8270.

8. Lorenz, D. A.; Song, J. M.; Garner, A. L., High-Throughput Platform Assay Technology for the Discovery of pre-microRNA-Selective Small Molecule Probes. *Bioconjugate Chem* **2015,** *26* (1), 19-23.

9. Lorenz, D. A.; Kaur, T.; Kerk, S. A.; Gallagher, E. E.; Sandoval, J.; Garner, A. L., Expansion of cat-ELCCA for the Discovery of Small Molecule Inhibitors of the Pre-let-7-Lin28 RNA-Protein Interaction. *Acs Med Chem Lett* **2018,** *9* (6), 517-521.

10. Song, J. M.; Menon, A.; Mitchell, D. C.; Johnson, O. T.; Garner, A. L., High-Throughput Chemical Probing of Full-Length Protein-Protein Interactions. *ACS Comb Sci* **2017,** *19* (12), 763-769.

11. Garner, A. L., cat-ELCCA: catalyzing drug discovery through click chemistry. *Chem. Commun.* **2018,** *54*, 6531-6539.

12. Sherman, E. J.; Lorenz, D. A.; Garner, A. L., Click Chemistry-Mediated Complementation Assay for RNA-Protein Interactions. *ACS Comb Sci* **2019,** *21* (7), 522-527.

13. Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V., NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chemical Biology* **2016,** *11* (2), 400-408.

14. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **2008,** *3* (6), 373-82.

15. Mili, S.; Steitz, J. A., Evidence for reassociation of RNA-binding proteins after cell lysis: Implications for the interpretation of immunoprecipitation analyses. *RNA* **2004,** *10* (11), 1692-1694.

16. Peabody, D. S., The RNA binding site of bacteriophage MS2 coat protein. *EMBO J* **1993,** *12* (2), 595-600.

17. Rackham, O.; Brown, C. M., Visualization of RNA–protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *The EMBO Journal* **2004,** *23* (16), 3346-3355.

18. D'Agostino, V. G.; Sighel, D.; Zucal, C.; Bonomo, I.; Micaelli, M.; Lolli, G.; Provenzani, A.; Quattrone, A.; Adami, V., Screening Approaches for Targeting Ribonucleoprotein Complexes: A New Dimension for Drug Discovery. *SLAS Discov* **2019,** *24* (3), 314-331.

19. Lorenz, M., Visualizing protein-RNA interactions inside cells by fluorescence resonance energy transfer. *RNA* **2009,** *15* (1), 97-103.

20. Huranova, M.; Jablonski, J. A.; Benda, A.; Hof, M.; Stanek, D.; Caputi, M., In vivo detection of RNA-binding protein interactions with cognate RNA sequences by fluorescence resonance energy transfer. *RNA* **2009,** *15* (11), 2063-2071.

21. Duan, N.; Arroyo, M.; Deng, W.; Cardoso, M. C.; Leonhardt, H., Visualization and characterization of RNA-protein interactions in living cells. *Nucleic Acids Res* **2021,** *49* (18), e107.

CHAPTER 3

A Live Cell Assay for the Detection of pre-miRNA-Protein Interactions[1]

As discussed in Chapter 1, due to the role of miRNA in regulating the expression of genes involved in virtually all cellular processes and the role of RBPs in regulation of miRNA processing, the interactions between miRNAs and their RBPs are prime candidates for therapeutic intervention and drug discovery. However, as summarized in Chapter 2, the platforms currently available for the detection of RPIs are technologically limited. To address this gap in technology, and building upon the advantages and learnings from the development of cat-ELCCA and SEA[1-6], a novel assay technology, RNA interaction with Protein-mediated Complementation Assay, or RiPCA, was designed and developed to enable detection of RPIs in live cells.

Inspired by SEA, RiPCA utilizes NanoBiT, a split luciferase reporter, to monitor the binding of an RNA and protein. Since RiPCA is a cellular assay, it precludes the use of click chemistry as a means of labeling the RNA-of-interest as in cat-ELCCA and, therefore, the engineered activity of HaloTag was leveraged to label the RNA probe with SmBiT. HaloTag, as previously discussed, is a lab-evolved dehalogenase that covalently binds to substrates containing chloroalkane handles.[7] This provides the advantage of permanently labeling the RNA-of-interest

[1] Reproduced in part from *RSC Chem. Biol.*, 2021, **2**, 241-247 with permission from the Royal Society of Chemistry. Additional details on the methods reported in this chapter were published in *Current Protocols*, 2022, **2**, e358. The work presented in this chapter was followed up on the assay conceptualized by Dr. Daniel A. Lorenz.

in the cell. Thinking forward to further engineering of this assay, it was posited that utilizing a protein-based labeling system could enable interaction detection in specific cellular compartments.

In RiPCA, cells stably expressing SmBiT fused to HaloTag (SmHT) are transiently co-transfected with (1) a plasmid encoding an RBP-of-interest fused to LgBiT and (2) an RNA probe

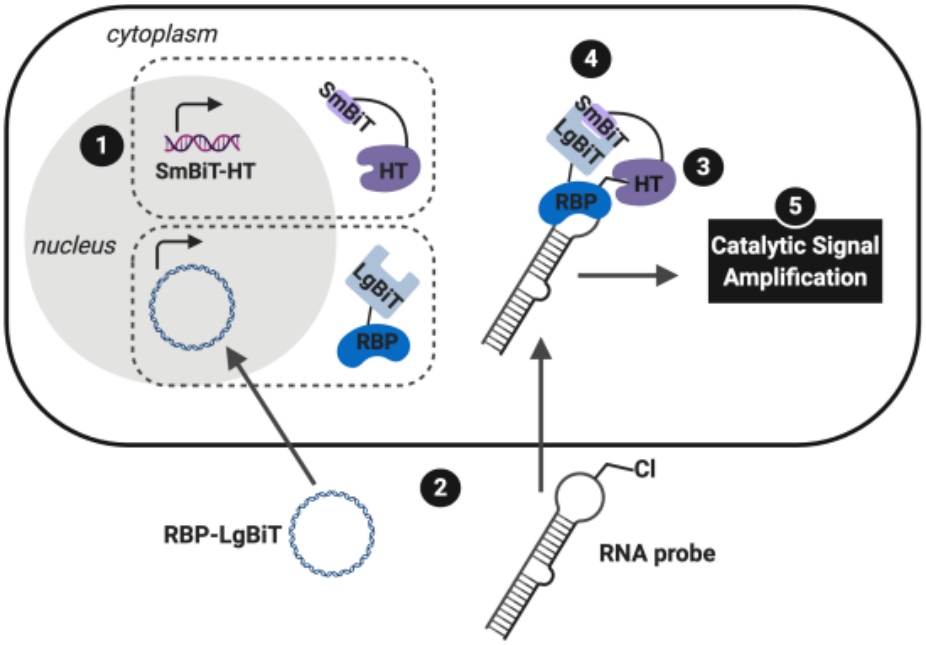


Figure 3.1 Schematic of RiPCA. Created with BioRender.com.

functionalized with a chloroalkane motif. In the cell, the RNA probe is covalently labeled with SmBiT via HaloTag and subsequent interaction between the RBP and RNA brings Sm- and Lg-BiT in close enough proximity to induce reconstitution of functional NanoLuc, enabling chemiluminescent detection of the interaction upon treatment with a NanoLuc substrate (Fig. 3.1).

There are four key advantages between RiPCA has over existing complementation assays. (1) RiPCA utilizes a relatively small chemical handle, avoiding the use of protein-binding RNA affinity tags (e.g., MS2 hairpins[8-12]), making it possible to probe RPIs involving non-mRNA species, including small or highly processed RNAs (e.g., microRNA). (2) Due to the weak intrinsic affinity between SmBiT and LgBiT (K_d of 190 μM), signal generation is driven by the RPI.[13] (3)

The use of a chemiluminescent readout yields enhanced sensitivity and favorable assay statistics relative to fluorescence-based technologies. (4) Interaction dynamics could be readily monitored due to the reversible nature of the SmBiT/LgBiT interaction.

3.1 Optimization of RiPCA with Lin28 as Proof-of-Concept

As in previous assays, let-7/Lin28 was used as a model system to facilitate the optimization of RiPCA. In RiPCA there are three key elements that are required for successful signal generation: SmHT, a LgBiT-tagged RBP, and an RNA probe. First, to eliminate the need for triple transfection, FlpIn HEK 293 cells stably expressing SmHT were generated. The SmHT element was chosen to be stably expressed rather than the RBP-Lg to allow for greater modularity in future iterations of the assay. Stable expression of SmHT was confirmed by confocal microscopy (Fig. 3.2). Next, plasmids encoding Lin28A with the LgBiT fusion at either the N- or the C-terminus were generated. The synthetic RNA probes, repurposed from Dicer cat-ELCCA[3, 14], contain a 5' biotin handle and an aminoallyluridine base located in the terminal loop that enables conjugation to a

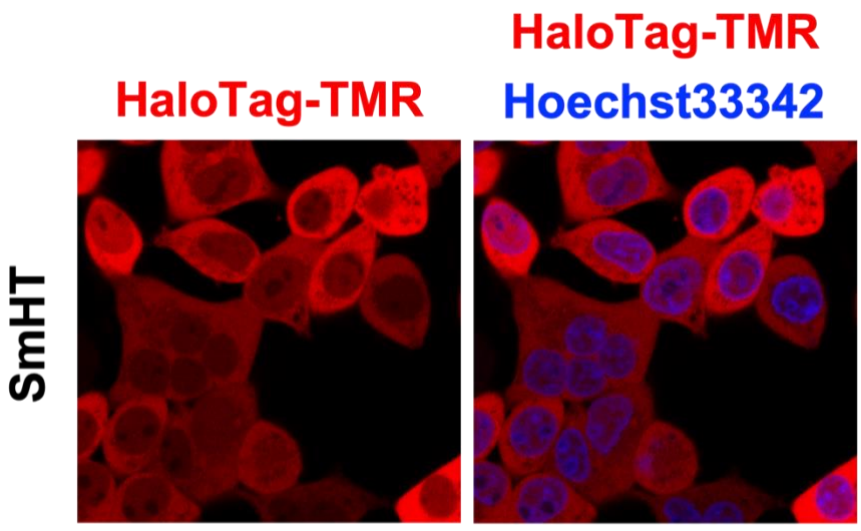


Figure 3.2. Confirmation of SmBiT-HaloTag (SmHT) expression in Flp-In HEK 293 cells via confocal microscopy. SmHT (red) was detected following conjugation with a HT ligand labeled with tetramethylrhodamine (TMR). Nuclei were visualized with Hoechst 3334.

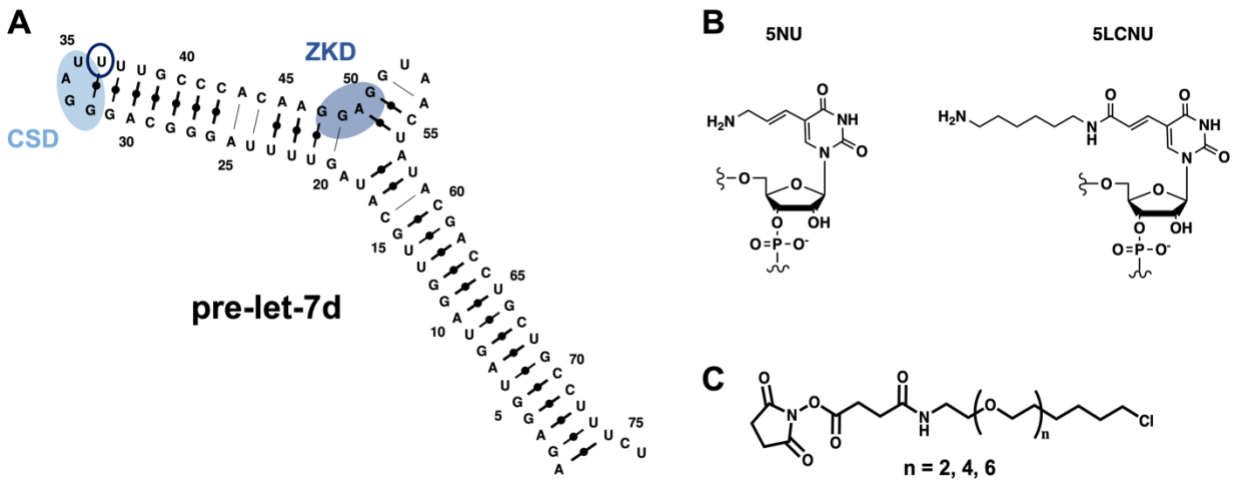


Figure 3.3. Design of RiPCA RNA probe. (A) Sequence of pre-let-7d probe. Lin28 domain binding sites are highlighted. The cold shock domain (CSD) binding site is highlighted in light blue and the zinc knuckle domain (ZKD) binding site is highlighted in dark blue. Location of the modified uridine residue is circled. **(B)** Structures of the modified bases 5-aminoallyl uridine (5NU) (left) and 5-aminonexylacrylamino uridine (5LNCU) (right). **(C)** Structure of HaloTag ligands.

HaloTag ligand via N-Hydroxysuccinimide (NHS) coupling. Due to changes in manufacturer supply, some probes, namely pre-miR-21 and pre-miR-34a, are modified with 5-aminoallyluridine (5NU) while others with 5-aminohexylacrylamino uridine (5LCNU) (Fig. 3.3B). As exemplified by data collected with pre-miR-21, there was no significant difference in the signal generated with the pre-miR-21 probes containing the different modified uridines, nor was there any difference when the 5LCNU-modified pre-miR-21 probe was conjugated to HaloTag ligands with various length PEG linkers (Fig. 3.5B).

As expected, RiPCA cells transfected with let-7d-Cl and either Lin28A-LgBiT or LgBiT-Lin28A generated substantial chemiluminescent signal (Fig. 3.4A). To validate that RiPCA signal was a result specifically of the formation of a complex between let-7d-Cl/SmHT and LgBiT-tagged Lin28A, RiPCA cells were transfected with two negative control combinations of RNA and protein: let-7d-Cl/LgBiT and let-7d-TCO (not reactive with HT)/Lin28A-LgBiT (Fig. 3.3A). The negative control transfections produced minimal signal compared to the positive control transfections of let-7d-Cl with either Lin28A-LgBiT or LgBiT-Lin28A (Fig. 3.4A). Since there

was little difference between the signal generated by Lin28A-LgBiT or LgBiT-Lin28A, only the Lin28A-LgBiT construct was utilized in future optimization experiments. To further demonstrate that signal is specifically generated upon interaction between the SmBiT-tagged RNA and LgBiT-tagged protein, cells were co-transfected with pre-miR-21 or pre-let-7d and increasing amounts of the corresponding unlabeled probe. While competition between labeled and unlabeled pre-miR-21

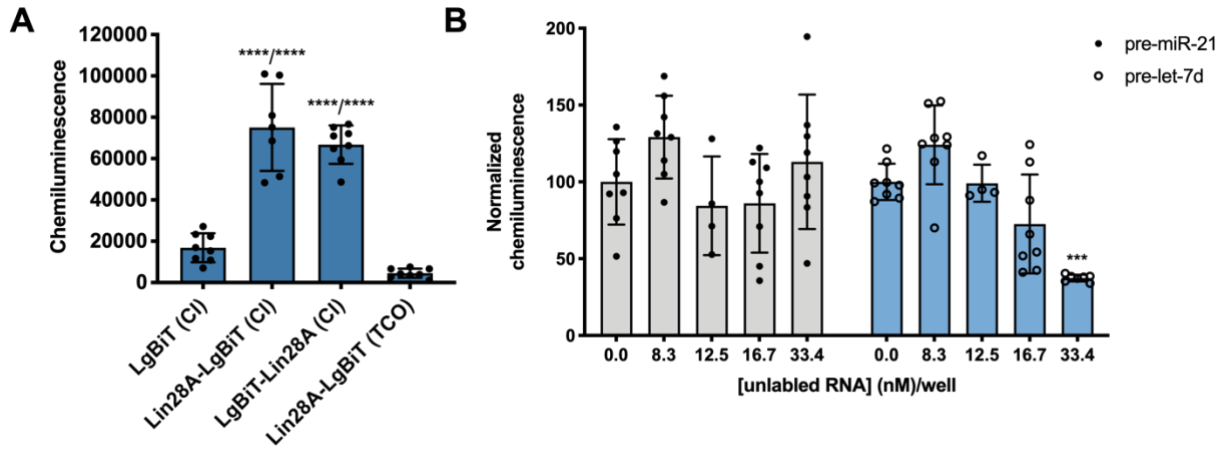


Figure 3.4. RiPCA proof-of-concept. (A) Selective detection of the Lin28A/pre-let-7d interaction in RiPCA. **(B)** Selective inhibition of pre-let-7d signal by competition with increasing concentrations of unlabeled pre-miRNA probe. Chemiluminescence signal was normalized to the signal generated by the wells transfected only with HT ligand-labeled pre-miRNA. See Tables 3.1 and 3.2 for *p* values.

did not inhibit RiPCA signal, unlabeled pre-let-7d did decrease signal generated in a dose-dependent manner (Fig. 3.4B).

Next, several modifications to the let-7d probe were tested to explore optimal probe design. RiPCA was performed with probes that differed at the 5' position to confirm that the 5' biotin modification did not inhibit let-7/Lin28 interaction in RiPCA. In fact, signal-to-background (S/B) defining pre-miR-21 readout as the background signal, herein the main metric used to assess quality of RiPCA data, was greater with the biotinylated probes, possibly due to increased stability of the 5' biotinylated RNA probes (Fig. 3.5A). Moving forward with the biotinylated probes, the effect of changing the length of the PEG linker between the NHS ester and chlorine handle of the HaloTag ligand was measured. As the linker length increased from PEG2 to PEG4 and PEG6,

there was a slight decrease in S/B observed (Fig. 3.5C). Given the consistency of S/B with PEG4 probes, that HaloTag ligand was used throughout future experiments in this chapter. Lastly, there was no notable difference in S/B when RiPCA was performed with RNA probes that contained a free amine modification at the 5' position rather than in the terminal loop of the pre-miRNA hairpin to measure effects of changing the site of the SmHT labeling (Fig. 3.5D). Given that Lin28 binds

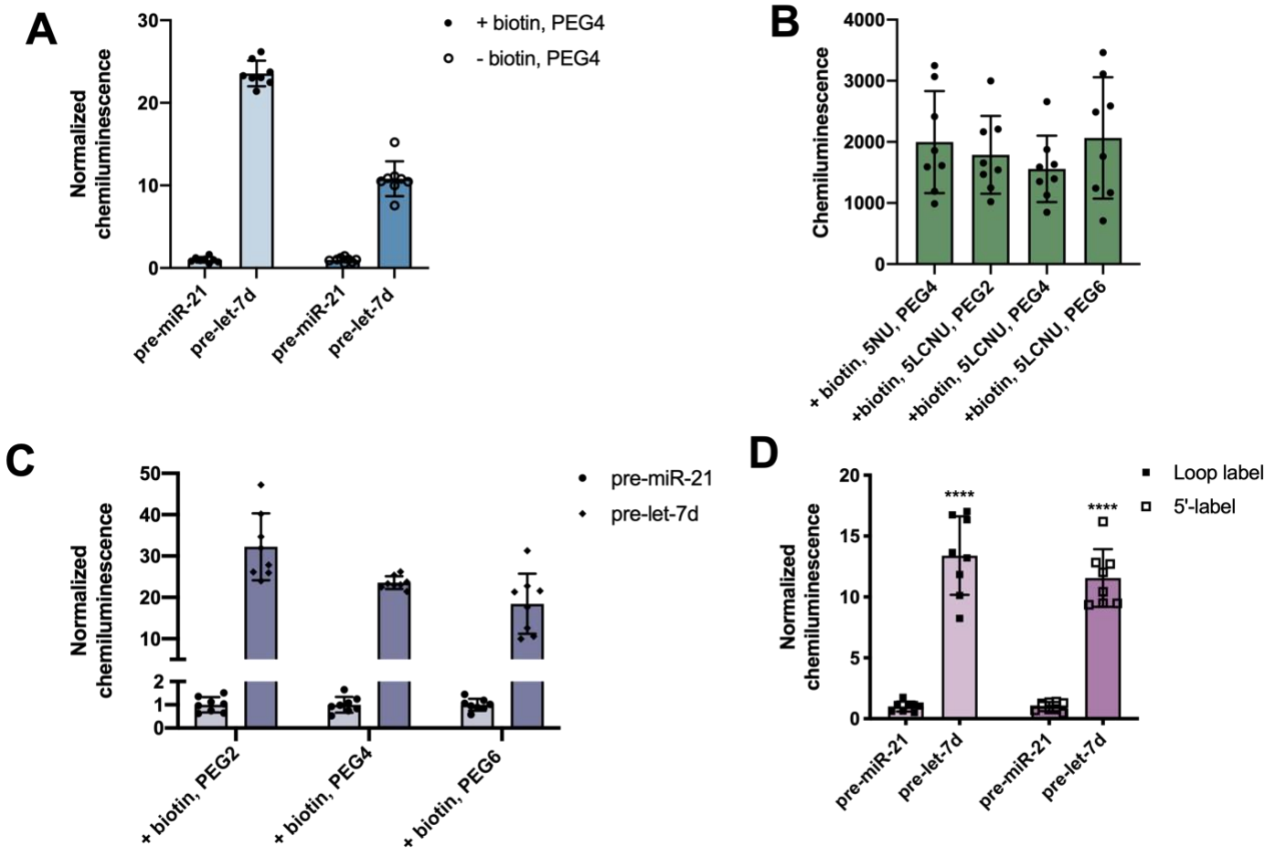


Figure 3.5. Pre-miRNA probe optimization. Chemiluminescence signal detection of the interaction between Li28A-LgBiT and **(A)** pre-miR-21 and pre-let-7d probes with or without a 5' biotin modification, **(B)** pre-miR-21 probes containing various HaloTag ligand PEG linker lengths and either 5-aminoallyluridine (5NU) or 5-aminohexylacrylamino uridine (5LCNU), **(C)** pre-miR-21 and pre-let-7d probes conjugated to HaloTag ligands with varying PEG linker lengths, and **(D)** pre-miRNA probes containing the site of HaloTag ligand conjugation in either the terminal loop or the 5' end of the hairpin. Normalized chemiluminescence is reported as the signal produced by pre-let-7d for each condition divided by the signal produced by pre-miR-21 for the corresponding condition. See Table 3.3 for *p* values.

with pre-let-7 through interactions with motifs in the terminal loop, the loop label RNA probes were used in subsequent experiments.

To assess the dynamic range of RiPCA, the effect of increasing the amount of DNA or RNA transfected per well was measured. Increasing the amount of DNA transfected per well while keeping the amount of RNA transfected constant resulted in a dose dependent decrease in S/B, a trend which could be attributed to higher background signal generated by elevated levels of cellular Lin28A-LgBiT allowing for increased non-specific interactions with SmBiT (Fig. 3.6A). When the amount of RNA transfected per well relative to DNA was increased, there was a corresponding increase in S/B (Fig. 3.6B), demonstrating that RiPCA detects the specific interaction between let-7/Lin28. Further, these data provide evidence that, in this setup, the chloroalkane-modified RNA probe is the limiting reagent for signal generation. Based on the consistency of the data, and keeping in mind the cost of reagents, transfecting 1 ng DNA and 8.3 nM RNA per well were determined to be the optimal conditions.

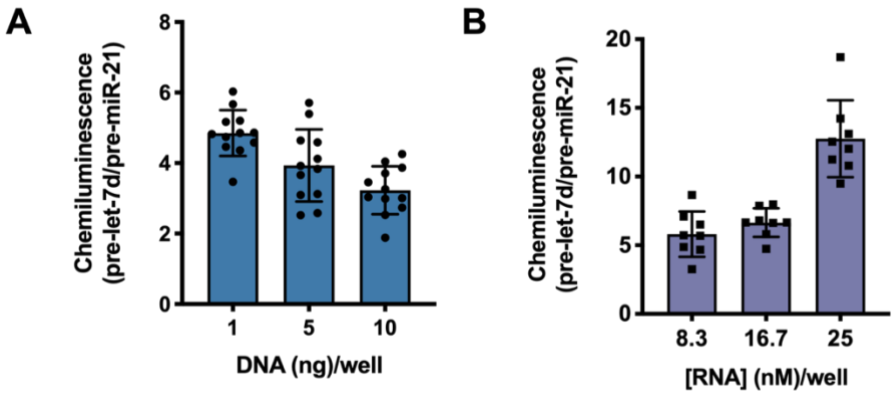


Figure 3.6. DNA and RNA dependence in RiPCA. (A) Dependence of S/B on the amount of Lin28A-LgBiT plasmid and **(B)** pre-miRNA-Cl probe transfected in SmHT-expressing cells.

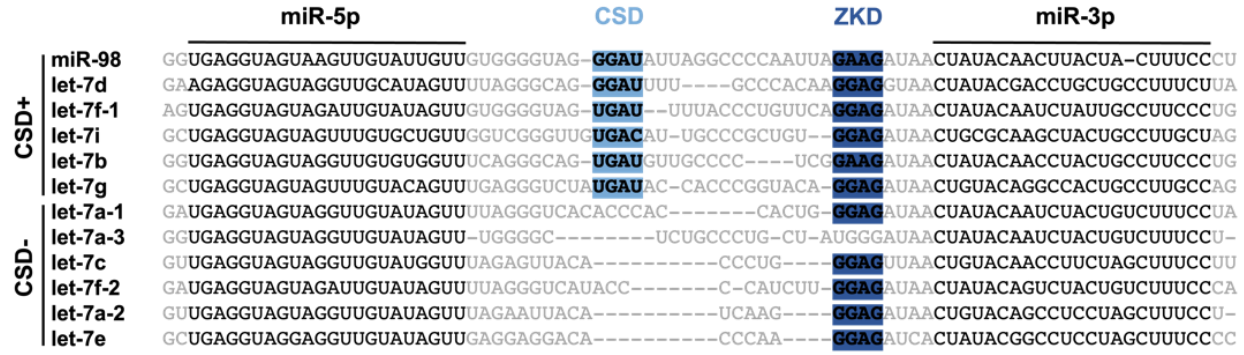


Figure 3.7. Let-7 family of miRNA. Alignment of pre-let-7 hairpins with mature 5p and 3p miRNA sequences are marked in black. The CSD and ZKD binding sites are highlighted in light blue and dark blue, respectively. Sequences containing both CSD and ZKD binding sites are labeled CSD+ and those containing only the ZKD binding site are labeled CSD-. Figure adapted from *Molecular Cell*, **2018**, 71, 271-283.

3.2 Specificity and sensitivity of Lin28 RiPCA: domains & Lin2

As introduced in Chapter 1, Lin28 binds the terminal loop of pre-let-7 via the CSD and ZKD, which bind GGAU and GGAG motifs, respectively (Fig. 3.7). Biochemical studies of the individual Lin28 RBDs revealed that while the ZKD has greater sequence specificity than the CSD, the CSD has greater affinity for pre-let-7 than the ZKD and disproportionately influences the binding affinity of full-length Lin28A.[15-17] In a 2018 study, Ustianenko *et al.* found that Lin28 displayed preferential binding for the pre-let-7 isoforms that contain both the CSD and ZKD binding sites (CSD^+) relative to sequences that lack the CSD binding motif (CSD^-).[18] CSD^+ sequences were similarly enriched with Lin28 in a proteomics study to a greater degree than CSD^- sequences.[19]

First, to determine whether the assay could differentiate between binding and non-binding sequences for Lin28A, RiPCA was performed with several pre-miRNA probes (Fig. 3.8A). As expected, very low signal was detected for pre-miR-34a, a non-binding sequence, above the pre-miR-21 background signal (Fig. 3.8B). Additionally, greater S/B was observed for the two CSD^+ sequences probed, pre-let-7d and -7g, relative to the CSD^- sequence probed, pre-let-7a-1 (Fig.

3.8B, Lin28A). In line with previous reports, in RiPCA, Lin28A demonstrated higher affinity for pre-let-7d relative to pre-let-7g.[18] Similarly, when RiPCA was performed with a CSD-LgBiT construct, greater S/B was observed with the CSD$^+$ sequences (Fig. 3.8B, CSD). As hypothesized due to the lower sequence specificity of the ZKD, there was much less differentiation between binding to the various pre-let-7s seen with the ZKD-LgBiT construct (Fig. 3.8B, ZKD). Unsurprisingly, Lin28 has a much lower K_d for pre-let-7 than the individual domains. When evaluated *in vitro*, Lin28A bound to pre-let-7g with a K_d of ~50 nM, which was ~4- ~15-fold lower than the K_ds for the CSD and ZKD, respectively.[15-16] Consistent with this data, the relative binding

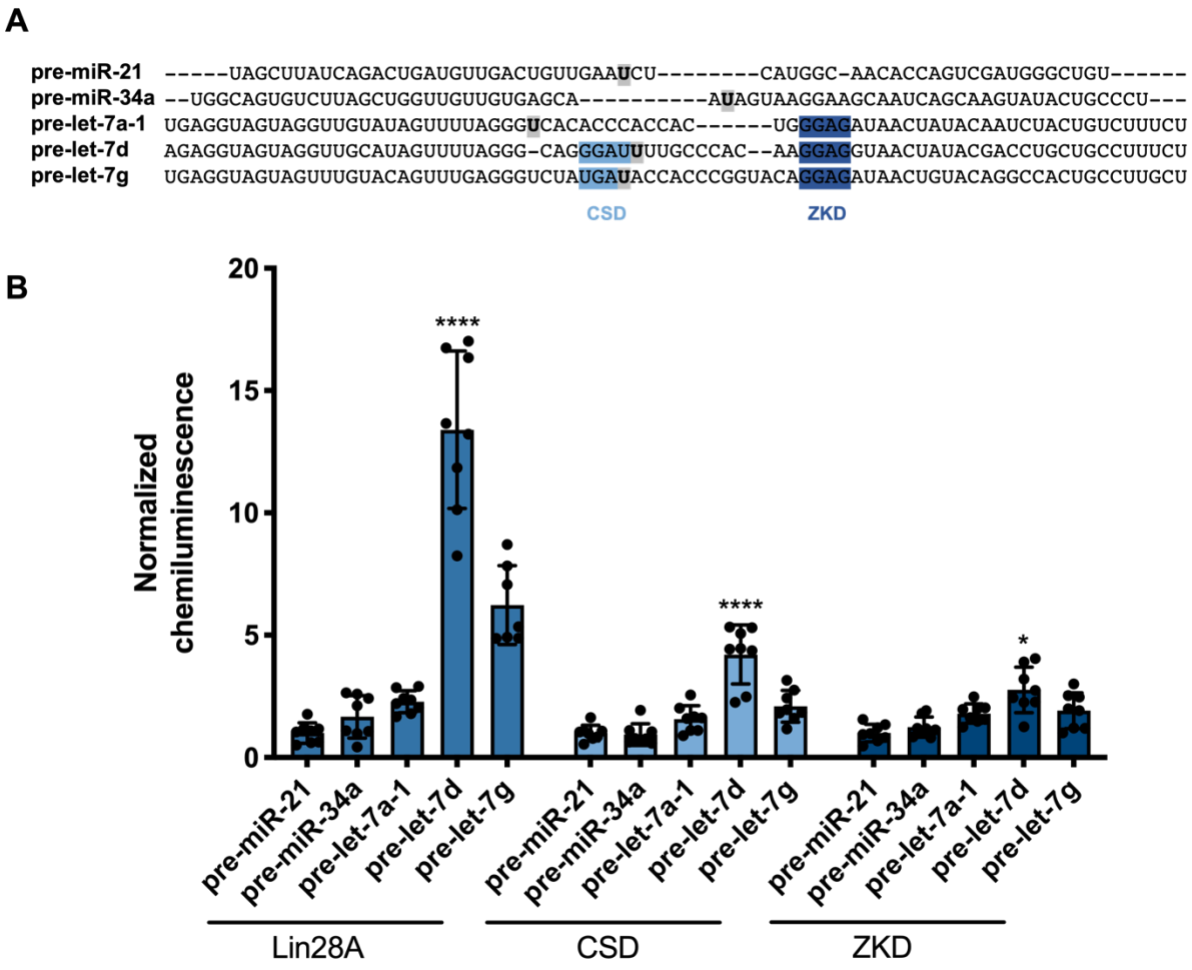


Figure 3.8. Selectivity of Lin28A in RiPCA. (A) Sequences of pre-miRNA probes used in RiPCA. The Lin28 CSD and ZKD binding sites are highlighted in light blue and dark blue, respectively. The location of the modified uridine is indicated by a bolded U. **(B)** Selectivity of Lin28A-LgBiT, CSD-LgBiT, and ZKD-LgBiT in RiPCA. See Table 3.4 for *p* values.

affinities remained consistent with Lin28A and the individual domains, but the S/B was significantly lower for CSD and lower still for ZKD, likely indicative of an affinity threshold under these conditions in RiPCA.

Because Flp-In HEK 293 cells express Lin28B, not Lin28A, RiPCA was performed with Lin28B-LgBiT. Signal detection in this system confirmed that endogenous RBP expression is not necessarily inhibitory in RiPCA. Binding was observed with Lin28B and the CSD^+ sequences, but not the CSD^- nor the non-binding sequences (Fig. 3.9). In line with previous findings, there was a slight binding preference for pre-let-7d relative to pre-let-7g (Fig. 3.9).

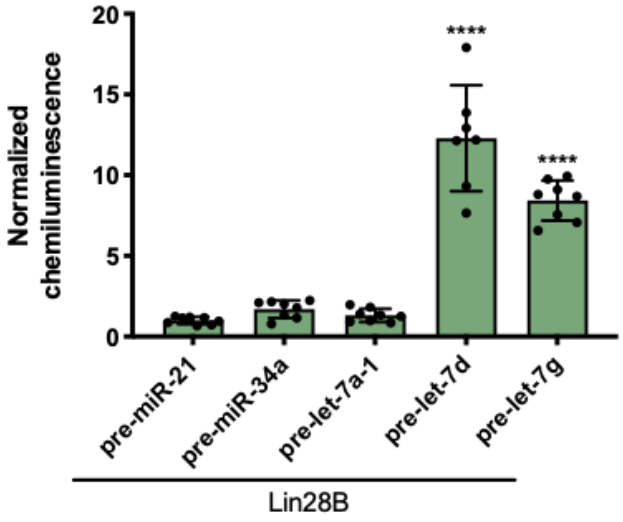


Figure 3.9. Selectivity of Lin28B in RiPCA. Selectivity of Lin28B-LgBiT against pre-miRNAs in the cytoplasm. See Table 3.5 for *p* values.

3.3 Nuclear RiPCA

Given that the subcellular localization of RNA and RPIs is tightly regulated and closely tied to their function,[20] it was appealing to assess whether RiPCA could be engineered to detect RPIs in distinct cellular compartments. Since many RPIs occur in the nucleus, a stable cell line expressing SmHT in the nucleus via a nuclear localization signal (NLS) was generated. Expression of SmHT-NLS in the nucleus was confirmed by confocal microscopy (Fig. 3.10). It has been shown that molecules <40 kDa are capable of entering the nucleus through the nuclear pore complex, therefore it was likely that the pre-miRNA-Cl probes could translocate to the nucleus.[21]

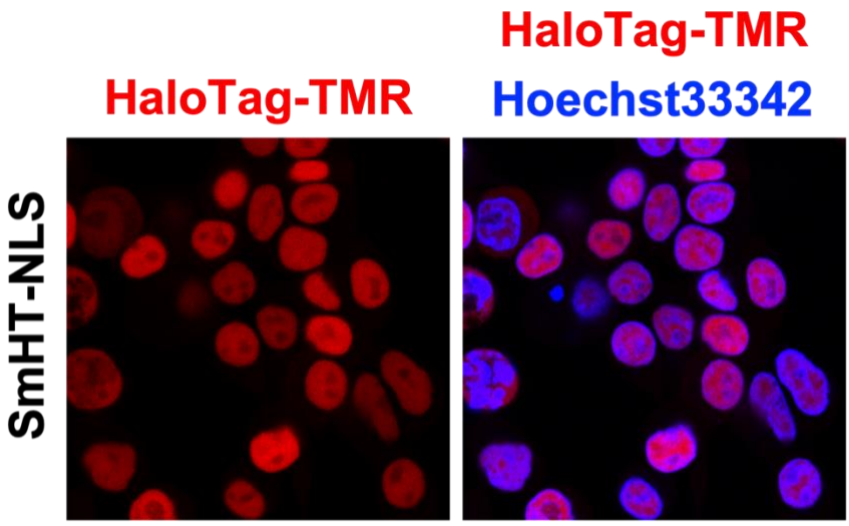


Figure 3.10. Confirmation of nuclear SmBiT-HaloTag (SmHT) expression in Flp-In HEK 293 cells via confocal microscopy. SmHT (red) was detected following conjugation with a HT ligand labeled with tetramethylrhodamine (TMR). Nuclei were visualized with Hoechst 33342 (blue).

As mentioned in Chapter 1, due to functional nuclear localization signals, Lin28B accumulates in the nucleus and inhibits let-7 biogenesis by the Microprocessor, and while Lin28A is primarily localized to and functions in the cytoplasm, there is evidence of some nuclear expression.[22-23] As seen in cytoplasmic RiPCA, both Lin28A-LgBiT and LgBiT-Lin28A generated greater signal when provided with pre-let-7d compared to pre-miR-21 (Fig. 3.11A). To similarly

probe the dynamic range in the nuclear assay, increasing amounts of DNA were transfected while the amount of RNA remained constant. Contrary to the trend observed in cytoplasmic RiPCA, increasing the amount of DNA transfected resulted in enhanced signal to background (Fig. 3.11B). Several factors, including reduced residence time of Lin28A-LgBiT in the nucleus, competition with endogenously expressed Lin28B, or lower uptake of the pre-miRNA probes in the nucleus, could contribute to the increased tolerance of LgBiT expression observed in nuclear RiPCA. Consistent with the observation in the cytoplasm that the amount of pre-miRNA-Cl probe defines the amount of tertiary complex formation, increasing the amount of RNA transfected per well, however, yielded a dose-dependent upward trend in S/B (Fig. 3.11C).

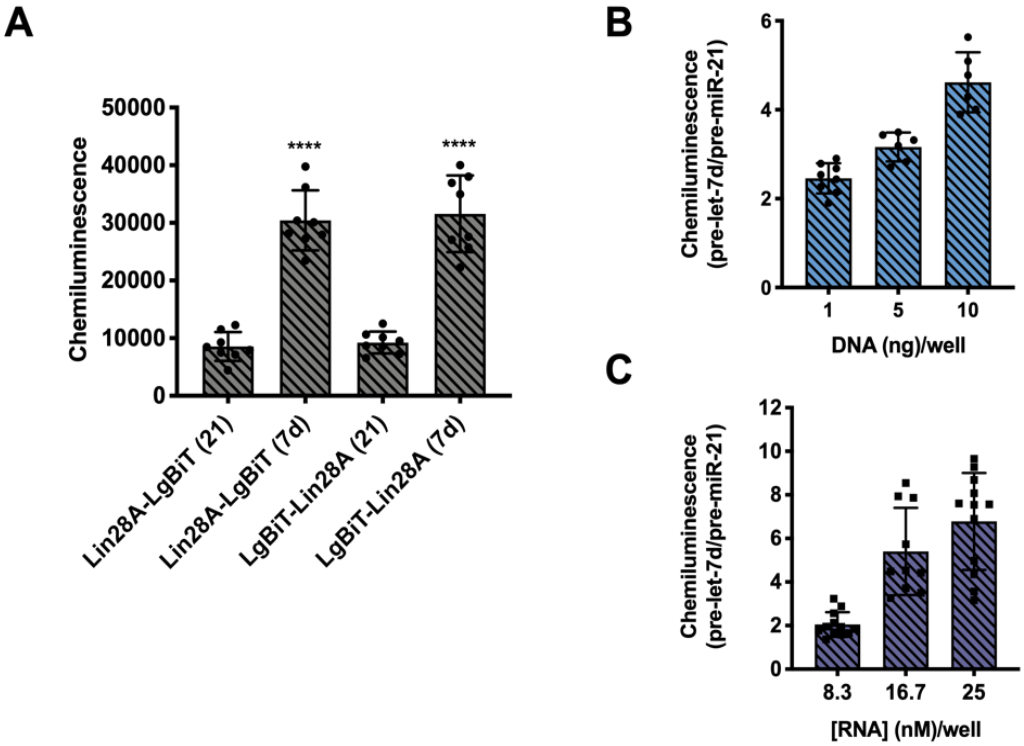


Figure 3.11. Nuclear RiPCA Proof-of-concept. (A) Selective detection of the Lin28A/pre-let-7d interaction in the nuclear RiPCA cell line expressing SmHT-NLS. Statistical significance determined by one-way ANOVA Sidak's multiple comparison test (n = 8); $p<0.0001$. **(B)** Dependence of S/B on the amount of Lin28A-LgBiT plasmid and **(C)** pre-miRNA-Cl probe transfected in the SmHT-NLS cell line.

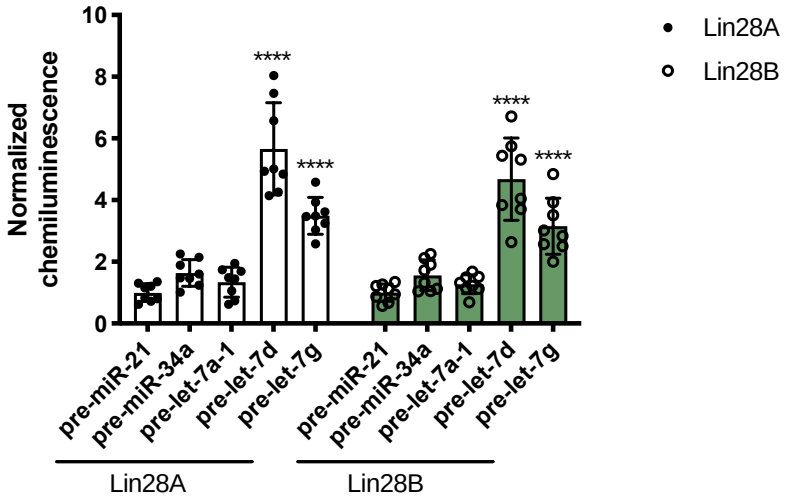


Figure 3.12. Nuclear RiPCA. Selectivity of Lin28A-LgBiT- and Lin28B-LgBiT against various per-miRNAs in the nuclear RiPCA. See Table 3.6 for *p* values.

Subsequently, RiPCA was performed with the various pre-miRNA probes and both Lin28A- and Lin28B-LgBiT in the nuclear SmBiT-HT cells. Nuclear RiPCA with both Lin28 isoforms generated S/B with CSD^+ sequences following the binding preference observed in cytoplasmic RiPCA (Fig. 3.12). Notably, lower S/B was observed in the nucleus, with the raw chemiluminescence signal being reduced 2.2- and 2.7-fold with Lin28A-LgBiT and LgBiT-Lin28A, respectively, in nuclear relative to cytoplasmic RiPCA (Fig. 3.8B and 3.11A). It is possible that the observed reduction in signal and S/B and higher tolerance for LgBiT expression in nuclear RiPCA could be attributed to reduced local expression of Lin28A-LgBiT and Lin28B-LgBiT or competition with endogenous Lin28B in the nucleus. Additionally, while molecules <40 kDa readily enter the nucleus through the nuclear pore complex[21], there could be reduced uptake of the pre-miRNA probes into the nucleus.

3.4 Conclusion

The development of RiPCA, a novel assay for the detection of cellular RPIs, provides a powerful tool for the validation and manipulation of RNA-protein interactions. RiPCA has been shown to selectively detect interactions between an RBP-of-interest and its pre-miRNA binding

partners. Through the development of the assay with Lin28, important points of assay optimization were identified and will be applied to the use of RiPCA with additional RPIs to be discussed in Chapter 4. Inhibition of signal by an unlabeled RNA probe demonstrates the potential for RiPCA to be utilized as a platform for the assessment and discovery of RPI inhibitors, which will be further discussed in Chapter 5. RiPCA is an extremely promising tool but is not without its limitations. Several implemented and proposed improvements to the assay protocol and design will be discussed in later chapters.

3.5 Methods

Materials:

Chemically synthesized pre-microRNAs (deprotected, desalted and HPLC purified), containing aminoallyl uridine or aminohexylacrylamino uridine modifications and biotin attached to the 5'-end of the sequence by an 18-atom spacer, were purchased from Dharmacon and used as received for the labeling reaction. HaloTag Succinimidyl Ester (O2) and (O4) Ligands were purchased from Promega and used as received (cat #1691 and #P6751). HaloTag Succinimidyl Ester containing PEG6 linker was synthesized following previously published protocol.[2] Note that the HaloTag Succinimidyl Ester Ligands should be dissolved and immediately portioned into single use aliquots stored at -80 °C to avoid degradation. Flp-In™-293 cells and associated vectors were purchased from ThermoFisher Scientific (Invitrogen cat #75007 and #601001, respectively). The Nano-Glo Live Cell Assay System was purchased from Promega and used as received (cat #N2012). HaloTag® TMR Ligand (Promega cat #G8251) and Lipofectamine™ RNAiMAX (Invitrogen cat #13778100) were used as received.

General cell culture methods:

Flp-In™-293 cells stably expressing either SmBiT-HaloTag or SmBiT-HaloTag-NLS were grown in DMEM (Corning cat #10-017-CV) supplemented with 10% FBS (Atlanta Biologicals S11550), L-glutamine (Gibco cat #25030081), and hygromycin B (100 μg/mL) (Gibco cat #10687010) at 37 °C with 5% CO_2 in a humidified incubator, passaged at least once before use for an experiment. Cells were passaged using Trypsin-EDTA (0.25%) (Gibco cat #25300054) approximately 10 times, and no more than 15 times, before returning to low passage stocks. To count cells, cells were harvested and 10 μL of the cell suspension was mixed with 10 μL Trypan Blue (Gibco cat #15250061) ([final] = 0.2% trypan blue) and counted using a hemocytometer.

General assay and data analysis methods:

Chemiluminescence data was collected on a BioTek Cytation3 plate reader. All data was analyzed using GraphPad Prism version 6.0c for Mac OS X (GraphPad Software, www.graphpad.com). All normalized chemiluminescence is reported as the signal of each well divided by the average signal of quadruplicate pre-miR-21 wells. The only exception is Fig. 2C, in which signal is normalized by dividing by the average of quadruplicate wells containing no unlabeled probe and multiplying by 100.

Statistical analysis:

All statistical tests were performed using Prism (v8). One- or two-way ANOVA tests were run for each set of data. Details of multiple comparisons tests are included in table legends. Graphs show mean ± standard deviation.

Table 3.1. Statistical significance associated with Fig. 3.4A. Statistical significance between chemiluminescence produced by cells co-transfected with chloroalkane-labeled pre-let-7d and Lin28A-LgBiT or LgBiT-Lin28A and cells co-transfected with either a LgBiT or TCO-labeled pre-let-7d control (ns = not significant). Statistical significance determined by one-way ANOVA Sidak's multiple comparison test (n = 8).

Construct	LgBiT control	TCO control
Lin28A-LgBiT	<0.0001 (****)	<0.0001 (****)
LgBiT-Lin28A	<0.0001 (****)	<0.0001 (****)

Table 3.2. Statistical significance associated with Fig. 3.4B. Statistical significance between chemiluminescence produced by cells transfected with chloroalkane-labeled pre-miR-21 and pre-let-7d sequences and cells transfected with varying amounts of unlabeled pre-miR-21 and pre-let-7d in cytoplasmic RiPCA (ns = not significant). Statistical significance determined by one-way ANOVA Dunnett's multiple comparison test (n = 8, except 12.5, for which n = 4). One data point from the pre-let-7d [33.4] data set was identified as an outlier using the ROUT method where Q = 1% (Prism) and eliminated from the data set and statistical analysis.

[Unlabeled RNA] (nM)	pre-miR-21	pre-let-7d
8.3	0.2845 (ns)	0.5244 (ns)
12.5	0.9738 (ns)	>0.9999 (ns)
16.7	0.9554 (ns)	0.3660 (ns)
33.4	0.9690 (ns)	0.0004 (***)

Table 3.3. Statistical significance associated with Fig. 3.5D. Statistical significance between chemiluminescence produced by cells transfected with pre-miR-21 and pre-let-7d with chloroalkane labels in either the loop or at the 5'-end of the probe. Statistical significance determined by two-way ANOVA Sidak's multiple comparison test (n = 8).

Sequence		Statistical significance
#1	#2	
pre-miR-21 (5' label)	**pre-let-7d (5' label)**	<0.0001 (****)
pre-miR-21 (loop label)	**pre-let-7d (loop label)**	<0.0001 (****)
pre-miR-21 (5' label)	**pre-miR-21 (loop label)**	>0.9999 (ns)
pre-let-7d (5' label)	**pre-let-7d (loop label)**	0.3927 (ns)

Table 3.4. Statistical significance associated with Fig. 3.8B. Statistical significance between signal-to-background (S/B) produced by various pre-miRNA-Cl sequences and pre-miR-21 with each LgBiT fusion in cytoplasmic RiPCA (ns = not significant). Statistical significance determined by two-way ANOVA Tukey's multiple comparison test (n = 8).

Sequence	Lin28A	CSD	ZKD
pre-miR-34a	0.7456 (ns)	>0.9999 (ns)	0.9922 (ns)
pre-let-7a-1	0.1464 (ns)	0.8381 (ns)	0.5962 (ns)
pre-let-7d	<0.0001 (****)	<0.0001 (****)	0.0152 (*)
pre-let-7g	<0.0001 (****)	0.2826 (ns)	0.4645 (ns)

Table 3.5. Statistical significance associated with Fig. 3.9. Statistical significance between signal-to-background (S/B) produced by various pre-miRNA-Cl sequences and pre-miR-21 with Lin28B-LgBiT in cytoplasmic RiPCA (ns = not significant). Statistical significance determined by one-way ANOVA Sidak's multiple comparison test (n = 8).

Sequence	Lin28B
pre-miR-34a	0.8874 (ns)
pre-let-7a-1	0.9929 (ns)
pre-let-7d	<0.0001 (****)
pre-let-7g	<0.0001 (****)

Table 3.6. Statistical significance associated with Fig. 3.12. Statistical significance between signal-to-background (S/B) produced by various pre-miRNA-Cl sequences and pre-miR-21 with Lin28A- or Lin28B-LgBiT fusion in nuc-RiPCA (ns = not significant). Statistical significance determined by two-way ANOVA Tukey's multiple comparison test (n = 8).

Sequence	Lin28A	Lin28B
pre-miR-34a	0.8330 (ns)	0.9073 (ns)
pre-let-7a-1	0.9970 (ns)	0.9995 (ns)
pre-let-7d	<0.0001 (****)	<0.0001 (****)
pre-let-7g	<0.0001 (****)	<0.0001 (****)

Cloning:

All constructs were generated using standard molecular cloning techniques.[24]

Table 3.7 Primers for RiPCA constructs. The sequences of primers utilized to clone various constructs for RiPCA.

Construct	Primers	Vector
SmBiT-HT	F: GTACGGTACCGCCACCATGGTGACCGGCTACCGG R: CAGTGCGGCCGCTCACTATTAGTGGTGATGGTGATGATG	pcDNA5/FRT
SmBiT-HT -Blp	F: GTACGGTACCGCCACCATGGTGACCGGCTACCGG R: GACTGCGGCCGCTCATTAGCTCAGCCCACCGGAAATCTCCAGAGTAG	
SmBiT-HT-NLS	F: TGAGTGGAGGTGGTCCAAAGAAAAAGAGAAAAGTATGGC R: TCAGCCATACTTTTCTCTTTTTCTTTGGACCACCTCCAC	pcDNA3
mLin28A-LgBiT	F: GTACGGTACCGCCACCATGGCCTCGGTGTCCAACC R: CAGTGCGGCCGCTCATTAAACACTGTTGATGGTTACTCG	
LgBiT	F: GTAGGATCCGCCACCATGGTCTTCACACTC R: CATCTCGAGACTGTTGATGGTTACTC	
LgBiT-mLin28A	F: ATGCCTCGAGATGGGCTCGGTGTCCAACC R: CATCTAGAATTCTGGGCTTCTGGGAGCA	
Lin28B-LgBiT	F: ATGGTACCGCCACCATGGCCGAAGGCG R: ATGCGATCGCTGTCTTTTTCCT	
mCSD-LgBiT	F: ATGGTACCGCCACCATGGCAGCGGAGAAGGCG R: TTGCGATCGCTCCTTTGGATCTTCG	
mZKD-LgBiT	F: ATGGTACCGCCACCATGAGTGAGCGGCGGCCAA R:AAGCGATCGCGCCCTGCTGGGCCT	

Preparation of RNA-HaloTag Ligand Conjugate

Amino-modified pre-miRNA (1.0 mM in 100 mM phosphate buffer, pH 8.0) was mixed with an equivalent volume of HaloTag ligand (10 mM in DMSO for O2 and O4; 20 mM in DMSO for O6). The reaction was then allowed to proceed at 25 °C for 1 h. pre-miRNA-Cl was precipitated by the addition of 0.11× volume of 3.0 M sodium acetate (pH 5.2) and 4 volume equivalents of

**Table 3.9. Sequence and modifications of terminal loop-labeled pre-miRNA probes. ***18 atom spacer is composed of hexaethylene glycol.

Sequence	5' modification	Sequence (CSD binding site bolded, ZKD binding site underlined, site of modified U bolded and underlined)	Length	Uridine modification
pre-miR-21	Biotin (18-atom spacer)*	UAGCUUAUCAGACUGAUGUUGACUGUUGAA**U**CUCAUGGCAACACCAGUCGAUGGGCUGUC	61	5NU
pre-miR-34a	Biotin (18-atom spacer)*	UGGCAGUGUCUUAGCUGGUUGUUGUGAGCAA**U**AGUAAGGAAGCAAUCAGCAAGUAUACUGCCCUA	65	5NU
pre-let-7a-1	Biotin (18-atom spacer)*	UGAGGUAGUAGGUUGUAUAGUUUUAGGG**U**CACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUCU	73	5LCNU
pre-let-7d	Biotin (18-atom spacer)*	AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAG**GGAU**UUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCU	76	5LCNU
pre-let-7g	Biotin (18-atom spacer)*	UGAGGUAGUAGUUUGUACAGUUUGAGGGUCUA**UGAU**ACCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCCUUGCU	79	5LCNU

Table 3.8. Sequence and modifications of 5'-labeled pre-miRNA probes.

Sequence	5' modification	Sequence (CSD binding site bolded, ZKD binding site underlined)	Length	Uridine modification
pre-miR-21	5' amine	UAGCUUAUCAGACUGAUGUUGACUGUUGAA UCUCAUGGCAACACCAGUCGAUGGGCUGUC	61	none
pre-let-7d	5' amine	AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAG**GGAU**UUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCU	76	none

cold ethanol, and pelleted at 20,000 × g for 40 min at 4 °C. The pellet was then re-suspended in 100 mM phosphate buffer (pH 8.0) at a concentration of 1.0 mM and stored at -80 °C.

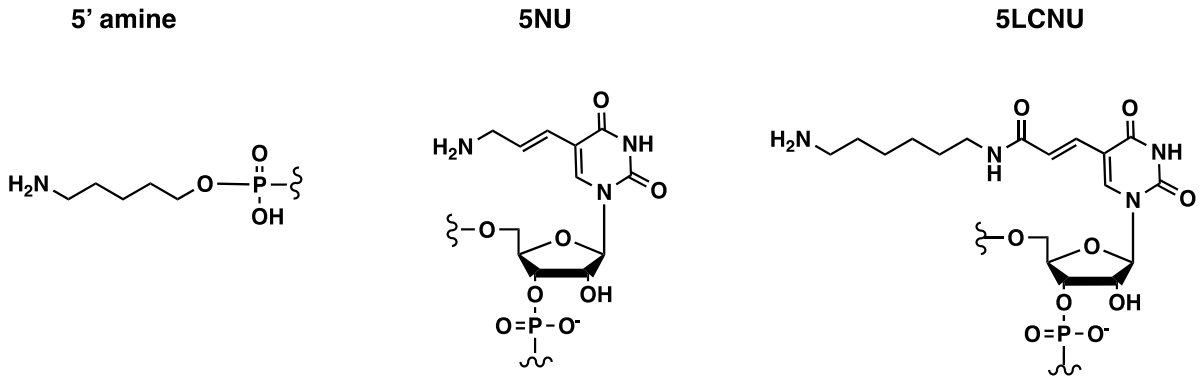


Figure 3.13. Structures of modified uridines. The structures of various amino modified uridines. From left to right, 5' amine, 5-aminoallyl uridine (5NU), and 5-aminohexylacrylamino uridine (5LCNU).

Generation of SmBiT-HT Stable Cell Lines

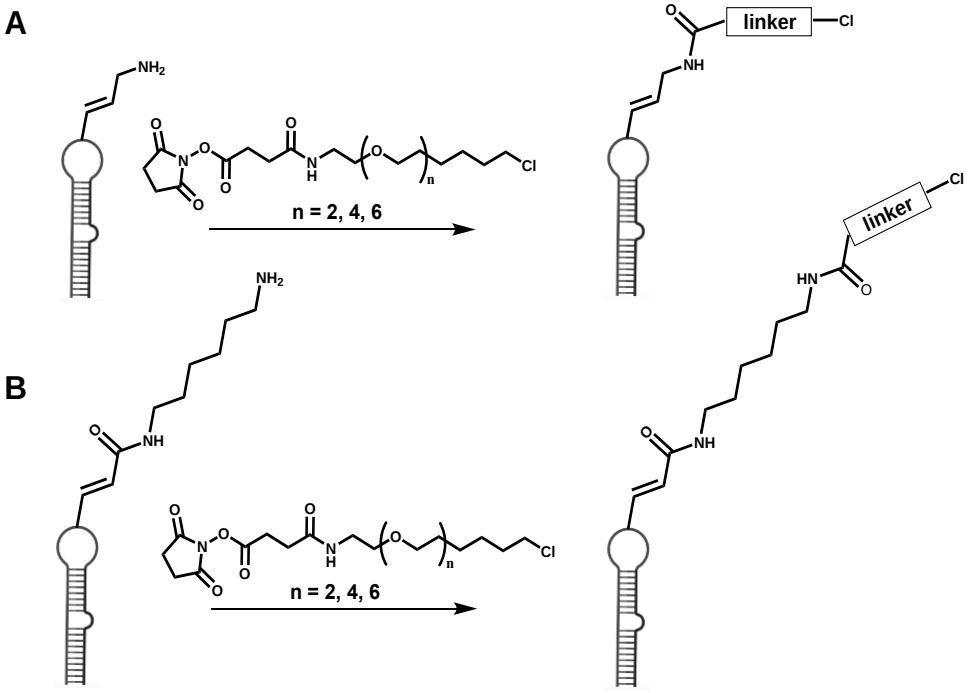


Figure 3.14. RNA labeling scheme. The structures of **(A)** 5-aminoallyl uridine and **(B)** 5-aminohexylacrylamino uridine modifications conjugated to HaloTag Succinimidyl ester ligands.

Flp-In-293 cells stably expressing a SmBiT-HT were generated by co-transfecting Flp-In-293 cells with 9 μg pOG44 and 1 μg pcDNA5/FRT using Lipofectamine™ LTX+PLUS reagent according to manufacturer's protocol. Expression and localization of SmBiT-HT or SmBiT-HT-NLS were confirmed by Western blot and confocal microscopy.

General RiPCA Protocol

Flp-In-293 cells stably expressing a SmBiT-HT protein were reverse transfected using Lipofectamine™ RNAiMAX Transfection Reagent. Cells were passaged approximately 10 times, and no more than 15 times, before returning to low passage stocks. To test "n" number of conditions, Solution A was prepared by combining 50 × (n+1) μL of room temperature Opti-MEM and 2.4 × (n+1) μL plasmid encoding selected RBP-LgBiT fusion. Solution B was prepared by adding pre-miRNA-Cl and plasmid (final concentrations 0.3 μM and 0.195 ng/μL, respectively) to 50 μL Opti-MEM™ for each separate condition to be tested. Solution B was mixed with 50 μL of Solution A to yield Solution A+B, which was incubated for at least 15 min at room temperature while cells were harvested. Cells were harvested as and counted as described above. Harvested cells were used to prepare Solution C, which was composed of 400 μL of 200,000 cells/mL. Solution C was mixed with 50 μL of Solution A+B and plated 100 μL per well, four wells per condition, in a white-bottom, tissue culture-treated 96-well plate (Corning cat #3917). The plate was incubated in a humidified incubator (37 °C and 5% CO_2) for 24 h. After incubation, the media was removed and replaced with 100 μL room temperature Opti-MEM™ and treated with 25 μL NanoGlo Live Cell Reagent diluted 1:20 according to the manufacturer's recommendation. All chemiluminescence data was collected immediately on a BioTek Cytation3 plate reader.

Representative calculations based on an assay for n = 5 conditions:

Solution A: Prepared for n+1= 6

6 x 50 µL → 300 µL OptiMEM™

6 x 2.4 µL → 14.4 µL Lipofectamine™ RNAiMAX

Solution B:

50 µL OptiMEM™

2.5 µL 3.9 ng/µL RBP-LgBiT plasmid

0.3 µL 50 µM pre-miRNA-Cl

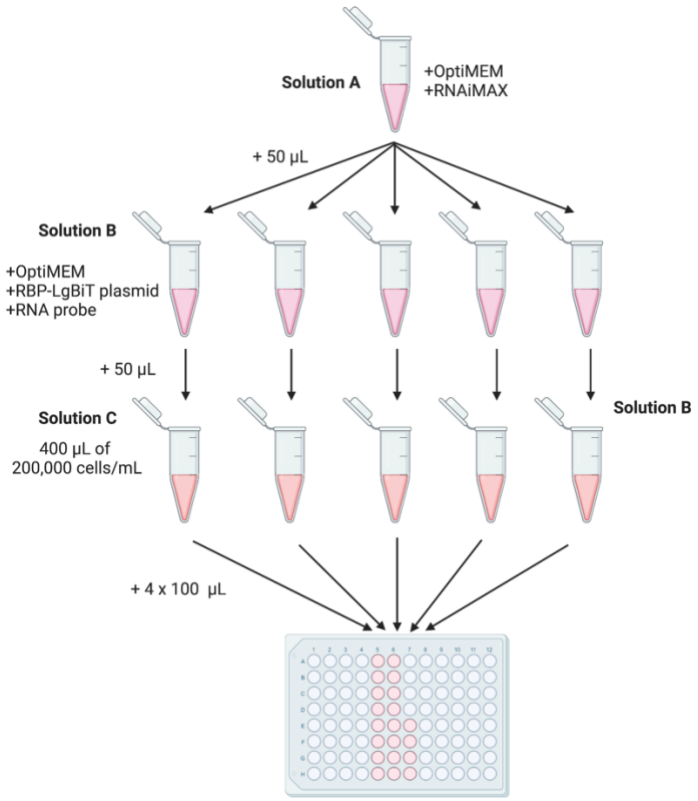


Figure 3.15. RiPCA transfection workflow.

DNA and RNA Titration RiPCA Protocol

To alter the amount of DNA transfected per well, higher concentration stocks of plasmid were used to allow addition of the same volume to each condition (final concentrations increased from 0.195 ng/μL to 0.39 ng/μL and 19.5 ng/μL). To alter the amount of RNA transfected per well, increasing volumes of 50 μM stock solution of pre-miRNA-Cl were used. The amount of Lipofectamine™ RNAiMAX was adjusted accordingly (from 2.4 × (n+1) μL for 16.7 nM/well to 1.2× (n+1) μL for 8.3 nM/well and 3.6 × (n+1) μL for 25 nM/well).

RiPCA Protocol for Competition with Unlabeled RNA

In competition experiments, the general RiPCA protocol was followed with the following change. In addition to the DNA and RNA added to Solution B, varying amounts of unlabeled pre-miRNA were added to Solution B (0, 0.15, 0.225, 0.3, or 0.6 μL unlabeled probe).

Confocal Microscopy

Flp-In cells stably expressing SmBiT-HT or SmBiT-HT-NLS were harvested and counted using methods described above. Cells were diluted to a density of 200,000 cells/mL and 200 μL was plated in an 8-well chambered coverglass (Nunc™ Lab-Tek™ II). The chambered coverglass was incubated in a tissue culture incubator (37 °C and 5% CO_2) for 24 h to allow the cells to adhere to the glass. To stain live cells, the media was supplemented with a single stain or a combination of stains at final concentrations of 50 nM HaloTag® TMR Ligand (Promega cat #G8251), 0.44 μM Hoescht 33342 (Fisher), and 0.2 μM MitoTracker Green FM (Cell Signaling Technology). The chambered coverglass was returned to the incubator for 30 min. The media was then removed and replaced with 200 μL Opti-MEM ™. Fluorescence was visualized using Nikon A1SI Confocal microscope. Images were processed with NIS-Elements.

3.6 Copyright

3.7 References

1. Lorenz, D. A.; Kaur, T.; Kerk, S. A.; Gallagher, E. E.; Sandoval, J.; Garner, A. L., Expansion of cat-ELCCA for the Discovery of Small Molecule Inhibitors of the Pre-let-7-Lin28 RNA-Protein Interaction. *Acs Med Chem Lett* **2018,** *9* (6), 517-521.

2. Song, J. M.; Menon, A.; Mitchell, D. C.; Johnson, O. T.; Garner, A. L., High-Throughput Chemical Probing of Full-Length Protein-Protein Interactions. *ACS Comb Sci* **2017,** *19* (12), 763-769.

3. Lorenz, D. A.; Garner, A. L., A click chemistry-based microRNA maturation assay optimized for high-throughput screening. *Chem. Commun.* **2016,** (52), 8267-8270.

4. Lorenz, D. A.; Garner, A. L., Approaches for the discovery of small molecule ligands targeting microRNAs. *Top. Med. Chem.* **2018,** *27*, 79-110.

5. Lorenz, D. A.; Vander Roest, S.; Larsen, M. J.; Garner, A. L., Development and implementation of an HTS-compatible assay for the discovery of selective small-molecule ligands for pre-microRNAs. *SLAS Disc.* **2018,** *23*, 47-54.

6. Sherman, E. J.; Lorenz, D. A.; Garner, A. L., Click Chemistry-Mediated Complementation Assay for RNA-Protein Interactions. *ACS Comb Sci* **2019,** *21* (7), 522-527.

7. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., HaloTag: a

novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **2008,** *3* (6), 373-82.

8. Lim, F.; Peabody, D. S., Mutations that increase the affinity of a translational repressor for RNA. *Nucleic Acid Res.* **1994,** *22*, 3748-3752.

9. Rackham, O.; Brown, C. M., Visualization of RNA–protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *The EMBO Journal* **2004,** *23* (16), 3346-3355.

10. Huranova, M.; Jablonski, J. A.; Benda, A.; Hof, M.; Stanek, D.; Caputi, M., In vivo detection of RNA-binding protein interactions with cognate RNA sequences by fluorescence resonance energy transfer. *RNA* **2009,** *15* (11), 2063-2071.

11. Bos, T. J.; Nussbacher, J. K.; Aigner, S.; Yeo, G. W., Tethered function assays as tools to elucidate the molecular roles of RNA-binding proteins. In *RNA Processing: Disease and Genome-wide Processing*, Yeo, G. W., Ed. Springer: Switzerland, 2016; Vol. 907, pp 61-88.

12. Graindorge, A.; Pinheiro, I.; Nawrocka, A.; Mallory, A. C.; Tsvetkov, P.; Gil, N.; Carolis, C.; Buchholz, F.; Ulitsky, I.; Heard, E.; Taipale, M.; Shkumatava, A., In-cell identification and measurement of RNA-protein interactions. *Nature Communications* **2019,** *10* (1), 5317.

13. Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V., NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chemical Biology* **2016,** *11* (2), 400-408.

14. Lorenz, D. A.; Song, J. M.; Garner, A. L., High-Throughput Platform Assay Technology for the Discovery of pre-microRNA-Selective Small Molecule Probes. *Bioconjugate Chem* **2015,** *26* (1), 19-23.

15. Nam, Y.; Chen, C.; Gregory, R. I.; Chou, J. J.; Sliz, P., Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **2011,** *147* (5), 1080-91.

16. Wang, L.; Nam, Y.; Lee, A. K.; Yu, C.; Roth, K.; Chen, C.; Ransey, E. M.; Sliz, P., LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. *Cell Reports* **2017,** *18* (11), 2664-2675.

17. Mayr, F.; Schutz, A.; Doge, N.; Heinemann, U., The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Res* **2012,** *40* (15), 7492-506.

18. Ustianenko, D.; Chiu, H.-S.; Treiber, T.; Weyn-Vanhentenryck, S. M.; Treiber, N.; Meister, G.; Sumazin, P.; Zhang, C., LIN28 selectively modulates a subclass of let-7 microRNAs. *Mol. Cell* **2018,** *71*, 271-283.

19. Treiber, T.; Treiber, N.; Plessmann, U.; Harlander, S.; Daiss, J. L.; Eichner, N.; Lehmann, G.; Schall, K.; Urlaub, H.; Meister, G., A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Molecular Cell* **2017,** *66* (2), 270-284.

20. Buxbaum, A. R.; Haimovich, G.; Singer, R. H., In the right place at the right time: visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* **2015,** *16* (2), 95-109.

21. D Angelo, M. A.; Hetzer, M. W., Structure, dynamics and function of nuclear pore complexes. *Trends Cell Biol.* **2008,** *10*, 456-466.

22. Piskounova, E.; Polytarchou, C.; Thornton, J. E.; LaPierre, R. J.; Pothoulakis, C.; Hagan, J. P.; Iliopoulos, D.; Gregory, R. I., Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* **2011,** *147* (5), 1066-79.

23. Heo, I.; Joo, C.; Cho, J.; Ha, M.; Han, J.; Kim, V. N., Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell* **2008,** *32* (2), 276-84.

24. Rosenblum, S. L.; Lorenz, D. A.; Garner, A. L., A live-cell assay for the detection of pre-microRNA–protein interactions. *RSC Chemical Biology* **2021,** (2), 241-247.

CHAPTER 4

Expansion of RiPCA for the Detection of Additional pre-miRNA-protein Interactions

In Chapter 1, several possible pre-miRNA binding proteins identified in a large-scale proteomics-based assay were introduced.[1] In particular, hnRNP A1 and Msi1/2 were highlighted as they are helpful examples of pre-miRNA-RBP interactions with some previous characterization.[2-5] Unlike the let-7/Lin28 RPI, there is much less biochemical and functional characterization of the interactions between the RBPs hnRNP A1 and Msi1/2 and their pre-miRNA binding partners. These interactions presented an opportunity to demonstrate the broad applicability of RiPCA to the detection of pre-miRNA-protein interactions and explore points of optimization in the assay, which will be discussed throughout this chapter.

4.1 Development of RiPCA 2.0

Prior to adapting RiPCA for detecting hnRNP A1- and Msi1/2-pre-miRNA interactions, it was noted that there was a significant amount of cell death occurring upon transfection with the Lipofectamine™ RNAiMAX reagent. A newly released transfection reagent, TransIT-X2® developed by Mirus, was identified as a promising alternative as it boasted lower cytotoxicity relative to another Lipofectamine™ reagent, Lipofectamine™ 2000.[6] Excitingly, in Lin28A RiPCA under very similar conditions (Lipofectamine™ RNAiMAX: 0.53 μL/well and X2: 0.33 μL/well), TransIT-X2® produced a ~2-fold greater average S/B in cytoplasmic RiPCA (15.5 vs. 30.1 with RNAiMAX and TransIT-X2®, respectively) and slightly greater S/B in nuclear RiPCA

(7.3 vs. 8.6 with RNAiMAX and TransIT-X2®, respectively) (Fig. 4.1A). TransIT-X2® was confirmed to be less toxic than RNAiMAX as it only resulted in 9.3% cell death in the cytoplasmic and 17.2% in the nuclear cell lines relative to the 45.6% and 41% cell death caused by transfection with RNAiMAX in the cytoplasmic and nuclear RiPCA cell lines (Fig. 4.1B). Based on the superior results and reduced cell death seen with TransIT-X2®, all subsequent experiments were performed with this transfection reagent and this iteration of the assay is hereafter referred to as RiPCA 2.0.

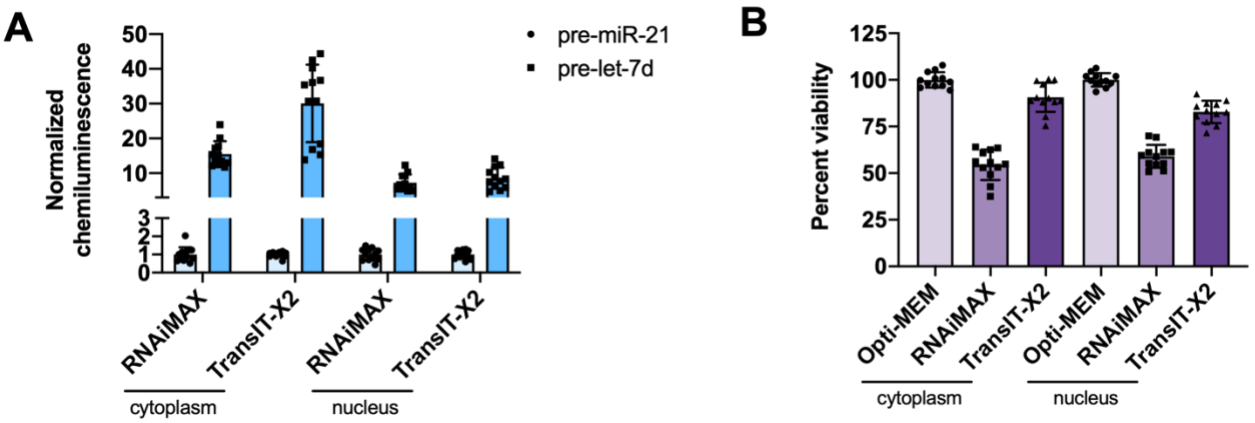


Figure 4.1. RNAiMAX vs. TransIT-X2®. (A) RiPCA with Lin28A-LgBiT and pre-miR-21 or pre-let-7d and either RNAiMAX or TransIT-X2® as the transfection reagent in the cytoplasm or the nucleus. **(B)** Cell viability measured by Cell Titer Glo of cytoplasmic or nuclear RiPCA cells transfected with Opti-MEM™ only (mock), RNAiMAX, or TransIT-X2®.

4.2 Optimization of RiPCA 2.0 for hnRNP A1 and Msi1/2

Of the pre-miRNA binding proteins identified by Treiber *et al.*, many shared similar binding profiles, and in particular, bound pre-let-7 miRNAs. Of note were hnRNP A1 and Msi1/2, which were highlighted in Chapter 1, and bind commonly to pre-let-7 hairpins, but also to pre-miR-18a.[1-4] Curious to assess whether RiPCA 2.0 could similarly define binding preferences within the let-7 family and detect interactions with another sequence, optimized assay conditions

for hnRNP A1, Msi1, and Msi2 were determined (Table 4.1). Due to the high signal produced by pre-let-7d in RiPCA with Lin28 and anticipated binding by hnRNP A1, Msi1, and Msi2, this sequence along with pre-miR-21 were utilized as the binding and non-binding sequences, respectively.

Using the previously identified conditions for Lin28 RiPCA 2.0, C-terminally tagged hnRNP A1, Msi1, and Msi2 constructs and pre-miRNA probes containing a PEG4 HaloTag ligand were used as a starting point for optimization. Starting with hnRNP A1, virtually no signal was detected when using the same amounts of DNA and RNA for transfection as used in Lin28 RiPCA 2.0 (Fig. 4.2A). When the amount of DNA was increased by 4-fold, modest S/B was detected with hnRNP A1-LgBiT (S/B of 4.2), but the signal produced by Msi1- and Msi2-LgBiT was only slightly above background (S/B of 2.6 and 1.9, respectively (Fig. 4.2B. Therefore, additional optimization was required.

The first point of optimization explored was the length of the PEG ligand in the RNA probes. The same assay conditions were repeated with RNA probes containing a PEG2 HaloTag ligand. Interestingly, these conditions yielded much higher S/B for each of the RBPs tested (S/B of 19.0, 18.9, and 34.9 for hnRNP A1-, Msi1-, and Msi2-LgBiT, respectively) (Fig. 4.2B). It is possible that, due to the long, flexible C-termini of the RBPs, the longer linker length positions the SmHT too far from LgBiT for successful enzyme reconstitution.

Next, the optimal position of the LgBiT tag on the RBP was determined. As described in Chapter 1, all three new RBPs tested in RiPCA 2.0 bind to RNA primarily through interactions with two N-terminal RRMs, whereas the C-terminus of each RBP is less structured. For all three new RBPs tested in RiPCA, higher S/B was measured when the LgBiT tag was appended to the C-terminus of the protein (Fig. 4.2B). N-terminal LgBiT-tagged constructs produced S/B that was

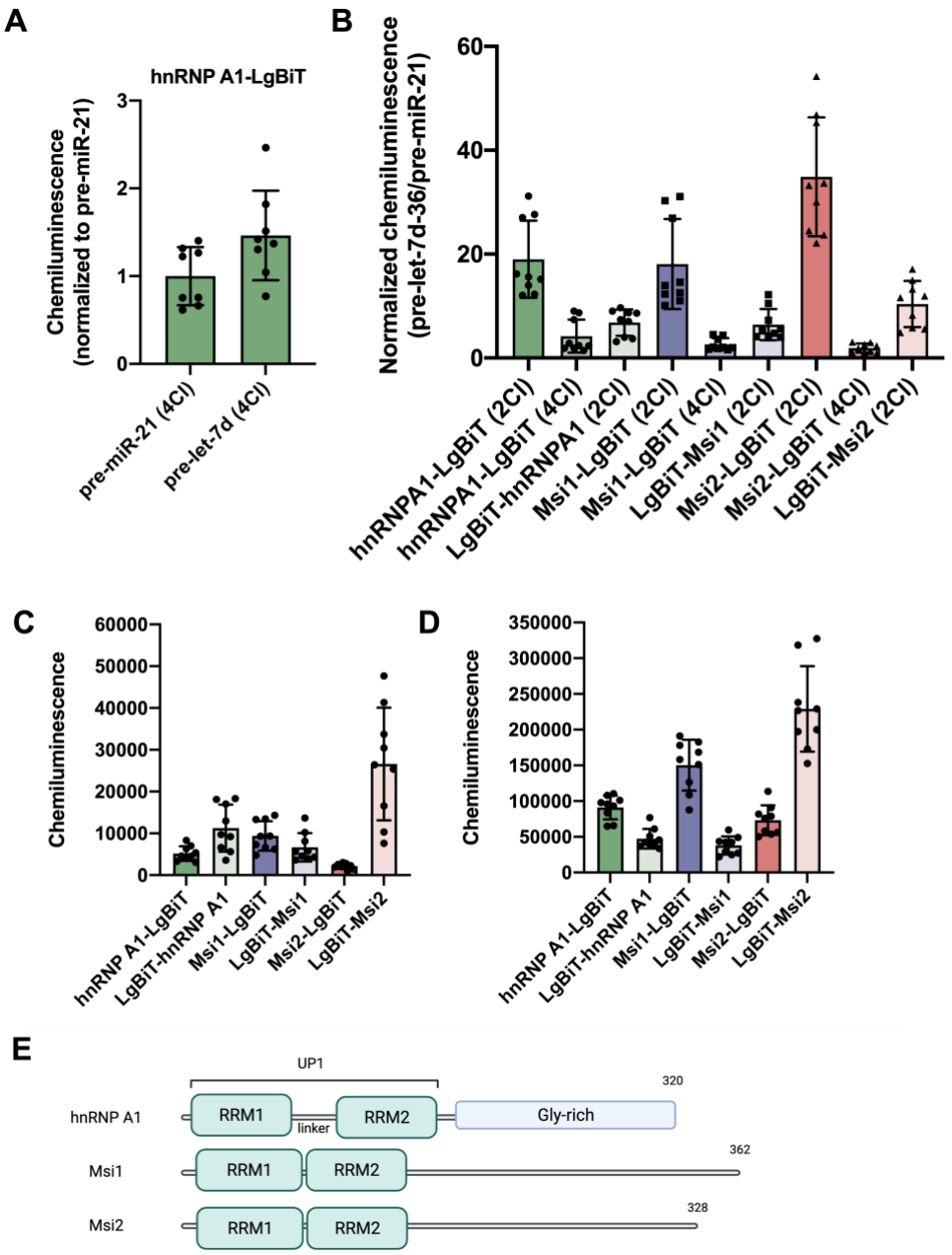


Figure 4.2. Optimization of RiPCA 2.0 for hnRNPA1, Msi1, and Msi2. Signal generated in RiPCA 2.0 with hnRNP A1-LgBiT **(B)** Determining the optimal LgBiT-orientation and linker length of the HaloTag ligand. Raw chemiluminescent signal generated by **(C)** pre-miR-21-2-Cl and **(D)** pre-let-7d RNA probes with N- or C-terminally LgBiT-tagged hnRNP A1, Msi1, and Msi2. **(E)** Domain maps of hnRNP A1, Msi1, and Msi2.

~3-fold lower for hnRNP A1, Msi1, and Msi2 (Fig. 4.2B). This differs from what was observed in Lin28 RiPCA, in which the N- and C-terminal LgBiT-tagged constructs produced similar S/B. Both Lin28A and Lin28B are smaller than hnRNP A1 and Msi1/2 by ~12-14 kDa and contain RBDs in both the N- and C-terminal domains. Given the common location of the RRMs in the N-

terminus of the RBPs and the unstructured C-terminal region (Fig 4.2E), it is possible that the N-terminal LgBiT tag hinders binding to the pre-miRNA probe and/or the flexibility of the C-termini of the RBPs enables optimal complex formation with the SmBiT-tagged pre-RNA probe.

To confirm expression of the constructs, a test expression in HEK 293 cells was performed and revealed that all constructs produced the LgBiT-tagged proteins to varying degrees (Fig. 4.3). It is important to note that the amount of DNA transfected to enable visualization by Western blot is significantly greater than the amount used in the assay. In Chapter 3, it was shown that in RiPCA, higher amounts of DNA, and consequently LgBiT-tagged RBP expression, results in decreased S/B due to higher background produced by nonspecific binding of Lg- and SmBiT. This is certainly the case with LgBiT-Msi2 as the background was, on average, ~12-fold greater with LgBiT-Msi2 than Msi2-LgBiT, whereas the signal produced was only an average of 3-fold greater with LgBiT-Msi2 (Fig. 4.2C, 4.2D, and 4.3). Given these data, the C-terminal LgBiT-tagged constructs were used to measure interactions in RiPCA 2.0.

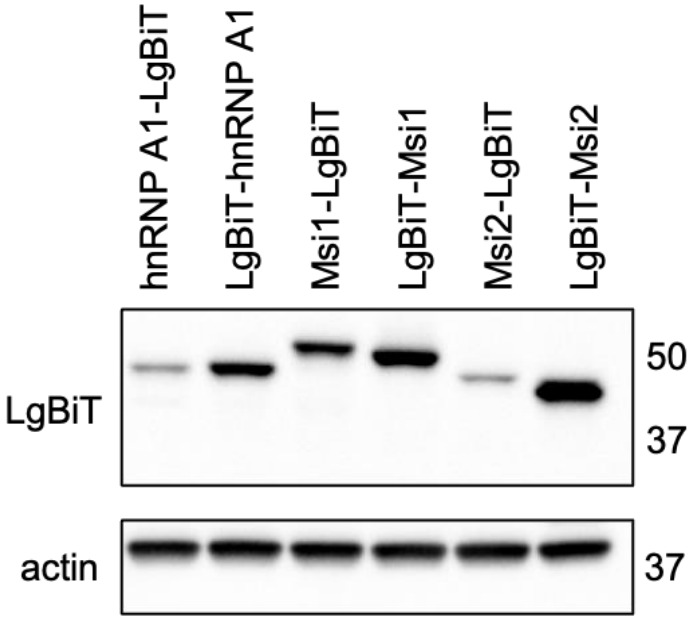


Figure 4.3. Expression of LgBiT plasmids. Expression of LgBiT-tagged proteins were visualized by Western Blot after transient transfection in HEK 293 cells.

4.3 Assessing Binding Preference with RiPCA 2.0

As with Lin28, the ability of RiPCA to discern relative affinities of hnRNP A1, Msi1, and Msi 2 was assessed. The previously prepared pre-miRNA probes, with the addition of another let-7 sequence, pre-miR-98, and a pre-miR-18a probe, were utilized to assess the binding preference of hnRNP A1-, Msi1, and Msi2-LgBiT and compare them to Lin28A- and Lin28B-LgBiT in RiPCA 2.0. Across all five RBPs tested, pre-let-7d produced the highest S/B of the five possible binding sequences (Fig. 4.4A-B). As expected, there was very little detectable interaction between

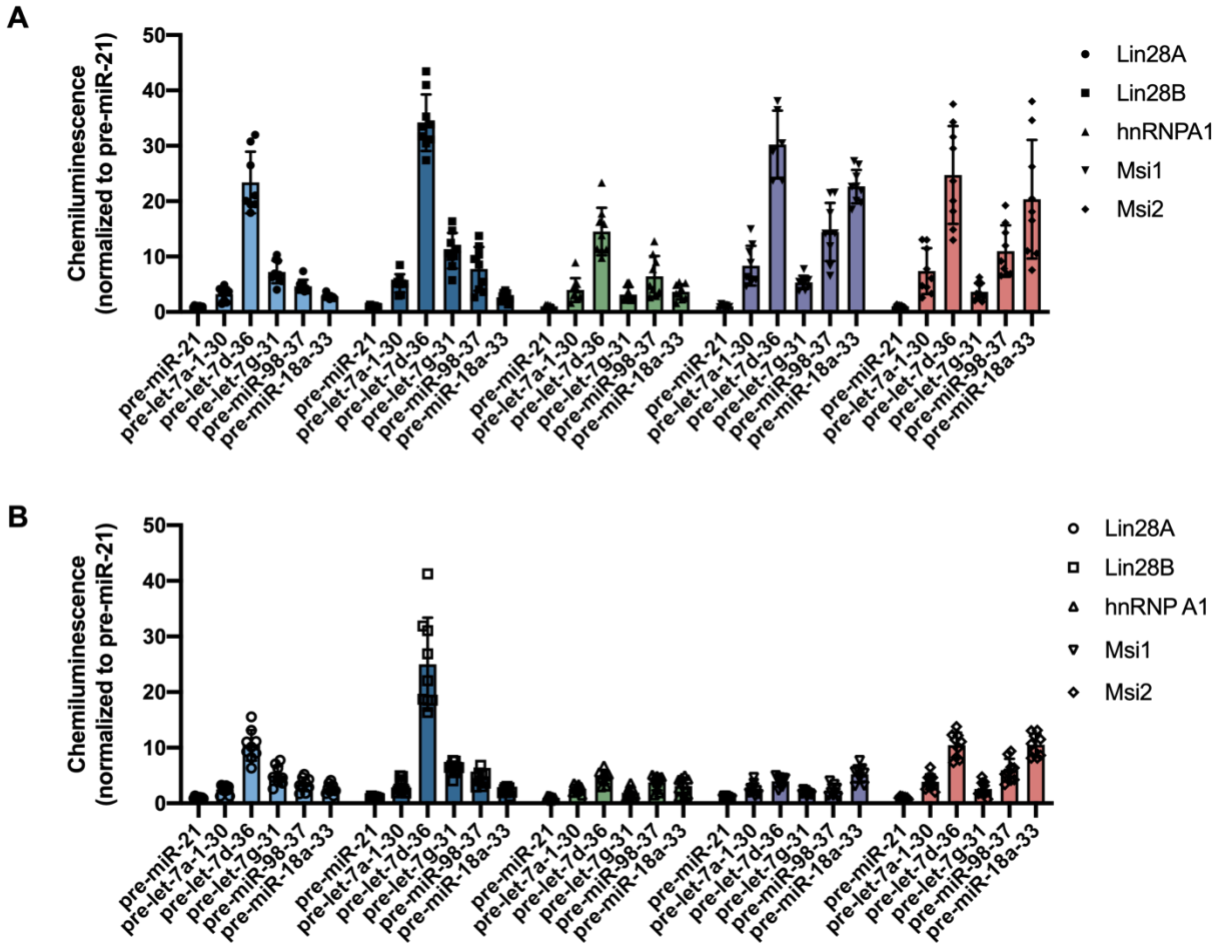


Figure 4.4. Profiling hnRNP A1, Msi1, and Msi2 binding preference in RiPCA 2.0. Signal generated in RiPCA 2.0 in the **(A) c**ytoplasm and **(B)** nucleus. Position of the uridine modification is noted as the final number in the probe label.

Lin28A- and Lin28B-LgBiT and pre-miR-18a and varying degrees of binding to the pre-let-7 sequences (Fig. 4.4A-B). Intriguingly, hnRNP A1-LgBiT showed much less preference for its two characterized binding partners, pre-let-7a1 and pre-miR-18a, relative to pre-let-7d and pre-miR-98 (Fig. 4.4A-B). The binding profile of Msi1- and Msi2-LgBiT mirror each other, but differ greatly from Lin28A, Lin28B, and hnRNP A-1 (Fig. 4.4A-B). Binding interactions between all but pre-let-7g and Msi1- and Msi2-LgBiT were detected, with greatest S/B being detected with pre-let-7d, pre-miR-98, and pre-miR-18a (Fig. 4.4A-B).

When tested in the nucleus, most RBPs produced much lower S/B. While Lin28A-, Lin28B-, and Msi2-LgBiT maintained consistent preferences among the sequences, the S/B produced was 2.3-, 1.4, and 2.4-fold lower than in the cytoplasm (Fig. 4.4A-B). There was a dramatic reduction in signal and specificity seen with both hnRNP A1- and Msi1-LgBiT (Fig. 4.4A-B). In the case of hnRNP A1, it is likely that this is due to endogenous competition with hnRNP A1, which is highly expressed in the nucleus.[7] However, Msi1 has been shown to primarily localize to the cytoplasm, which suggests that RiPCA is accurately detecting lower signal and S/B due to insufficient expression of Msi1-LgBiT in the nucleus.[8]

Next, a small library of pre-miRNA probes with different modification sites for HaloTag ligand conjugation was generated to further evaluate the ability of RiPCA 2.0 to validate site-specific binding to the pre-miRNA probes. All previously tested pre-miRNA probes contained the site for conjugation to the HaloTag ligand in the middle of the terminal loop, which were initially designed for Lin28 RPI detection and no further than 2 nt away from the CSD binding site (Fig. 4.5A). These modification sites are well within the reported hnRNP A1 footprint for pre-let-7a1 and pre-miR-18a, but the footprint covers a large portion of the terminal loop. Additionally, Msi1

and Msi2 bind UAG motifs and there are several throughout the selected pre-miRNA hairpins. Two additional probes for each of the binding pre-miRNA sequences were designed to contain the modification either closer to or within UAG motifs and closer to the beginning or end of, yet still within, the terminal loop (Fig. 4.5A).

First, it was confirmed in a gel shift assay, in which 500 nM pre-miRNA probe and 5 μM purified SmHT were incubated for 1 hour at room temperature, that all the pre-miRNA probes can react with SmHT (Fig. 4.6A and 4.6B). Next, all combinations of pre-miRNA probes and RBPs were assessed in both the cytoplasmic and the nuclear RiPCA cell lines (Fig. 4.5B and 4.5C). For

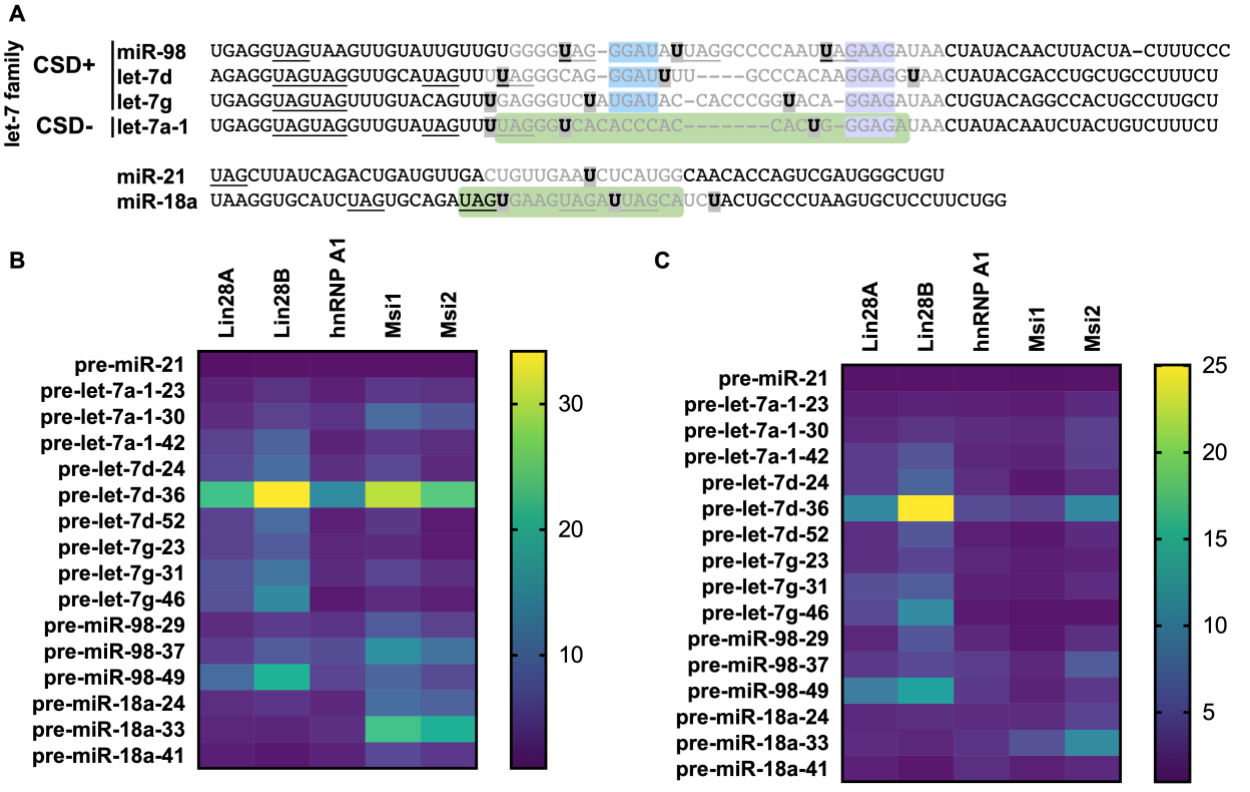


Figure 4.5. Exploring detection of site-specific pre-miRNA binding in RiPCA 2.0. (A) Pre-miRNA sequences used to generate RiPCA probes. The Lin28 CSD and ZKD binding sites are highlighted in light blue and lavender, respectively. The hnRNP A1 footprint as defined by REF is highlighted in light green. Msi1/2 UAG motifs are underlined. The sites of the modified uridines are highlighted in grey and bolded. Heat map of Lin28A-, Lin28B-, hnRNP A1-, Msi1-, and Msi2-LgBiT binding to the library of pre-miRNA probes in the **(B)** cytoplasm and **(C)** nucleus. Legend defines increasing signal-to-background (S/B) (signal divided by pre-miR-21) as color from blue to yellow. Position of the uridine modification is noted as the final number in the probe label.

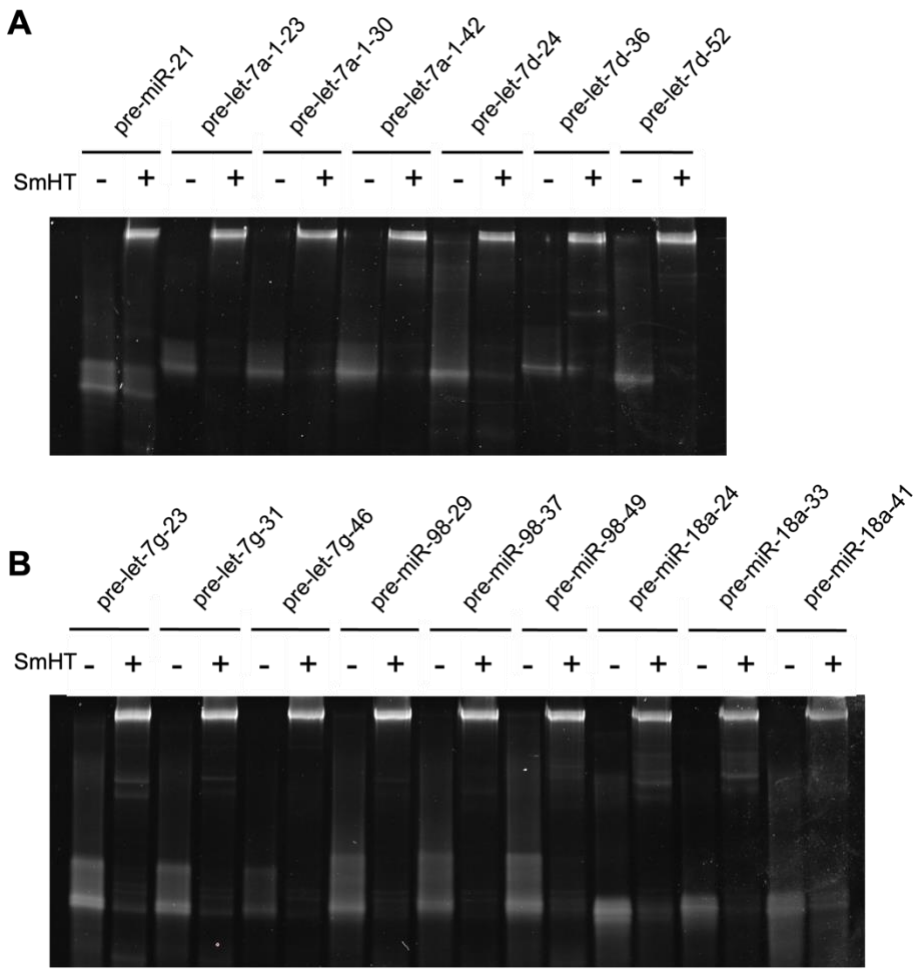


Figure 4.6. Gel shift assay with pre-miRNA probes and SmHT. Pre-miRNA probes were incubated with purified SmHT and run on **(A)** 15% or **(B)** 10% TBE-Urea gels next to RNA only samples. Position of the uridine modification is noted as the final number in the probe label.

most RBP/probe combinations, the greatest S/B was observed with the RNA probe containing the HaloTag ligand in the middle of the terminal loop. However, there were some notable exceptions. First, greater S/B was observed Lin28A- and Lin28B-LgBiT, but not the remainder of the RBPs, and pre-let-7a1-42, in which the modification site is located 2 nt from the ZKD binding site (Fig. 4.5A-C). Considering, pre-let-7a-1 is a CSD^- sequence, and thus is solely recognized via the Lin28 ZKD, this suggests the detection of site-specific binding in RiPCA 2.0 (Fig. 4.5A). Additional exceptions include Lin28A-LgBiT and pre-miR-98, as well as Lin28B-LgBiT with pre-let-7g and pre-miR-98, which produced slightly greater S/B with the probe containing the modification towards the 3' end of the terminal loop (Fig. 4.5B and 4.5C). This could be due to the resulting

conformation of the complex being optimally oriented for NanoLuc reconstitution. Interestingly, this phenomenon was not observed with pre-let-7d, in which the modification in pre-let-7d-52 is downstream of the ZKD binding site, whereas in pre-let-7g-46 and pre-miR-98-49, the modified uridine is upstream of the ZKD site (Fig. 4.5A-C).

Relative to the other four RBPs surveyed, hnRNP A1-LgBiT generated the lowest S/B and showed little-to-no sequence or site specificity in both cytoplasm and the nucleus (Fig. 4.5B and 4.5C). Interestingly, the two sequences which generated S/B with hnRNP A1-LgBiT, pre-let-7d-36 and pre-miR-98-37, possess the longest terminal loops of the probes tested, which could point to a requirement for RNA probe flexibility with this construct (Fig. 4.5A-C). As previously postulated, this could be due to competition with endogenously expressed hnRNP A1. Alternatively, this could be an indicator of lower affinity for the pre-miRNA sequences or shorter residence time relative to the other RBPs.

While there is a clear detection of binding preference for Msi1- and Msi2-LgBiT, there is less binding site specificity detected. Aside from the negative control, pre-miR-21, all pre-let-7g probes produced the lowest S/B with both Msi1- and Msi2-LgBiT, which is expected as the sequence does not contain UAG motifs in the terminal loop found in pre-let-7a-1, pre-let-7d, pre-miR-98, and pre-miR-18a. However, placing the uridine modification within or immediately adjacent to the UAG motif close to the 5’ end of the terminal loop did not enhance detection of binding to the probes (pre-let-7a-1-23, pre-let-7d-24, pre-miR-9829, and pre-miR-18a-24) (Fig. 4.5A-C). It is hypothesized that placing the uridine too close to the stem does not afford enough flexibility to form the requisite complex for signal detection.

Overall, these data suggest that, while it may be possible decipher site-specific binding preferences, perhaps the current system is not optimized to do so broadly for all RBPs. It is possible

that for some RPIs there may be a single, optimal position for the conjugation site in the RNA probe that is capable of generating successful RiPCA signal and S/B. While not available for some RPIs, structural data as well as RBP binding preference will be useful information when designing and testing RNA probes in RiPCA. Evidently, further exploration is required to fully elucidate the full capabilities and limitations of RiPCA.

4.4 Conclusion

The work presented in this chapter provides further evidence of the utility of RiPCA in detecting pre-miRNA-protein interactions. RiPCA has been successfully applied to the detection of the interactions between hnRNP A1, Msi1, and Msi2 and their respective pre-miRNA binding partners. The points of optimization required to detect S/B with these RBPs highlights the importance of determining optimal components, such as LgBiT orientation and HaloTag ligand linker length, and the modular nature of RiPCA. In exploring the sensitivity of RiPCA to the site of HaloTag ligand modification, it was found that the preferred site of modification is towards the middle of the terminal loop and there was limited site-specific binding information obtained and solely with Lin28A and Lin28B. Finally, the switch from Lipofectamine™ RNAiMAX to TransIT-X2® not only simplifies the transfection protocol but is less toxic to cells. Given RiPCA is intended to identify RPI inhibitors through loss of signal, the reduced cytotoxicity will be a particularly beneficial feature.

4.5 Methods

Materials:

Chemically synthesized pre-microRNAs (deprotected, desalted and HPLC purified), containing 5-aminohexylacrylamino uridine modifications and biotin attached to the 5'-end of the sequence by an 18-atom spacer, were purchased from Dharmacon and used as received for the labeling reaction. HaloTag Succinimidyl Ester (O2) and (O4) Ligands were purchased from Promega and used as received (cat #1691 and #P6751). Note that the HaloTag Succinimidyl Ester Ligands should be dissolved and immediately portioned into single use aliquots stored at -80 °C to avoid degradation. Flp-In™-293 cells and associated vectors were purchased from ThermoFisher Scientific (Invitrogen cat #75007 and #601001, respectively). The Nano-Glo Live Cell Assay System was purchased from Promega and used as received (cat #N2012). Transfection reagents Lipofectamine™ RNAiMAX (Invitrogen cat #13778100) and TransIT-X2® (Mirus cat #6000) were used as received. Cell Titer Glo was purchased from Promega and used as received (cat #G7570).

General cell culture methods:

Flp-In™-293 cells stably expressing either SmBiT-HaloTag or SmBiT-HaloTag-NLS were grown in DMEM (Corning cat #10-017-CV) supplemented with 10% FBS (Atlanta Biologicals S11550), L-glutamine (Gibco cat #25030081), and hygromycin B (100 μg/mL) (Gibco cat #10687010) at 37 °C with 5% CO_2 in a humidified incubator, passaged at least once before use for an experiment. Cells were passaged using Trypsin-EDTA (0.25%) (Gibco cat #25300054) approximately 10 times, and no more than 15 times, before returning to low passage stocks. To count cells, cells were harvested and 10μL of the cell suspension was mixed with 10μL Trypan Blue (Gibco cat #15250061) ([final] = 0.2% trypan blue) and counted using a hemocytometer.

General assay and data analysis methods:

Chemiluminescence data was collected on a BioTek Cytation3 plate reader. All data was analyzed using GraphPad Prism version 9.0.0 for Mac OS X (GraphPad Software, www.graphpad.com). All normalized chemiluminescence is reported as the signal of each well divided by the average signal of triplicate pre-miR-21 wells.

B. Cloning

hnRNPA1-LgBiT cloning. A synthetic human hnRNPA1 gene fragment was purchased from Twist Bioscience and inserted at the C-terminal position[9] using standard cloning techniques with KpnI and AsiSI restriction enzymes. Gene fragment contains a Kozak sequence at the N-terminus.

Gene Fragment:

5’GGTACCGCCACCATGTCTAAGTCAGAGTCTCCTAAAGAGCCCGAACAGCTGAGGA
AGCTCTTCATTGGAGGGTTGAGCTTTGAAACAACTGATGAGAGCCTGAGGAGCCATT
TTGAGCAATGGGGAACGCTCACGGACTGTGTGGTAATGAGAGATCCAAACACCAAG
CGCTCCAGGGGCTTTGGGTTTGTCACATATGCCACTGTGGAGGAGGTGGATGCAGCT
ATGAATGCAAGGCCACACAAGGTGGATGGAAGAGTTGTGGAACCAAAGAGAGCTGT
CTCCAGAGAAGATTCTCAAAGACCAGGTGCCCACTTAACTGTGAAAAAGATATTTGT
TGGTGGCATTAAAGAAGACACTGAAGAACATCACCTAAGAGATTATTTTGAACAGT
ATGGAAAAATTGAAGTGATTGAAATCATGACTGACCGAGGCAGTGGCAAGAAAAGG
GGCTTTGCCTTTGTAACCTTTGACGACCATGACTCCGTGGATAAGATTGTCATTCAG
AAATACCATACTGTGAATGGCCACAACTGTGAAGTTAGAAAAGCCCTGTCAAAGCA
AGAGATGGCTAGTGCTTCATCCAGCCAAAGAGGTCGAAGTGGTTCTGGAAACTTTG
GTGGTGGTCGTGGAGGTGGTTTCGGTGGGAATGACAACTTCGGTCGTGGAGGAAAC
TTCAGTGGTCGTGGTGGCTTTGGTGGCAGCCGTGGTGGTGGTGGATATGGTGGCAGT
GGGGATGGCTATAATGGATTTGGTAATGATGGAAGCAATTTTGGAGGTGGTGGAAG
CTACAATGATTTTGGGAATTACAACAATCAGTCTTCAAATTTTGGACCCATGAAGGG
AGGAAATTTTGGAGGCAGAAGCTCTGGCCCCTATGGCGGTGGAGGCCAATACTTTG
CAAAACCACGAAACCAAGGTGGCTATGGCGGTTCCAGCAGCAGCAGTAGCTATGGC
AGTGGCAGAAGATTTGCGATCGC

LgBiT-hnRNPA1 cloning. A synthetic human hnRNPA1 gene fragment was purchased from Twist Bioscience and inserted at the N-terminal position[9] using standard cloning techniques with XhoI and XbaI restriction enzymes.

Gene Fragment:

5’CTCGAGATGTCTAAGTCAGAGTCTCCTAAAGAGCCCGAACAGCTGAGGAAGCTCT
TCATTGGAGGGTTGAGCTTTGAAACAACTGATGAGAGCCTGAGGAGCCATTTTGAGC
AATGGGGAACGCTCACGGACTGTGTGGTAATGAGAGATCCAAACACCAAGCGCTCC
AGGGGCTTTGGGTTTGTCACATATGCCACTGTGGAGGAGGTGGATGCAGCTATGAAT
GCAAGGCCACACAAGGTGGATGGAAGAGTTGTGGAACCAAAGAGAGCTGTCTCCAG
AGAAGATTCTCAAAGACCAGGTGCCCACTTAACTGTGAAAAAGATATTTGTTGGTGG
CATTAAAGAAGACACTGAAGAACATCACCTAAGAGATTATTTTGAACAGTATGGAA
AAATTGAAGTGATTGAAATCATGACTGACCGAGGCAGTGGCAAGAAAAGGGGCTTT
GCCTTTGTAACCTTTGACGACCATGACTCCGTGGATAAGATTGTCATTCAGAAATAC
CATACTGTGAATGGCCACAACTGTGAAGTTAGAAAAGCCCTGTCAAAGCAAGAGAT
GGCTAGTGCTTCATCCAGCCAAAGAGGTCGAAGTGGTTCTGGAAACTTTGGTGGTGG
TCGTGGAGGTGGTTTCGGTGGGAATGACAACTTCGGTCGTGGAGGAAACTTCAGTG
GTCGTGGTGGCTTTGGTGGCAGCCGTGGTGGTGGTGGATATGGTGGCAGTGGGGAT
GGCTATAATGGATTTGGTAATGATGGAAGCAATTTTGGAGGTGGTGGAAGCTACAA
TGATTTTGGGAATTACAACAATCAGTCTTCAAATTTTGGACCCATGAAGGGAGGAAA
TTTTGGAGGCAGAAGCTCTGGCCCCTATGGCGGTGGAGGCCAATACTTTGCAAAACC
ACGAAACCAAGGTGGCTATGGCGGTTCCAGCAGCAGCAGTAGCTATGGCAGTGGCA
GAAGATTTTAATCTAGA

Msi1-LgBiT cloning. A synthetic human Musashi 1 (Msi1) gene fragment was purchased from Twist Bioscience and inserted at the C-terminal position[9] using standard cloning techniques with HindIII and AsiSI restriction enzymes. Gene fragment contains a Kozak sequence at the N-terminus.

Gene Fragment:

5’AAGCTTGCCACCATGGAGACTGACGCGCCCCAGCCCGGCCTCGCCTCCCCGGACT
CGCCGCACGACCCCTGCAAGATGTTCATCGGGGGACTCAGTTGGCAGACTACGCAG
GAAGGGCTGCGCGAATACTTCGGCCAGTTCGGGGAGGTGAAGGAGTGTCTGGTGAT
GCGGGACCCCCTGACCAAGAGATCCAGGGGTTTCGGCTTCGTCACTTTCATGGACCA
GGCGGGGGTGGATAAAGTGCTGGCGCAATCGCGGCACGAGCTCGACTCCAAAACAA
TTGACCCTAAGGTGGCCTTCCCTCGGCGAGCACAGCCCAAGATGGTGACTCGAACG

AAGAAGATCTTTGTGGGGGGGCTGTCGGTGAACACCACGGTGGAGGACGTGAAGCA
ATATTTTGAGCAGTTTGGGAAGGTGGACGACGCCATGCTGATGTTTGACAAAACCAC
CAACCGGCACCGAGGGTTCGGGTTTGTCACGTTTGAGAGTGAGGACATCGTGGAGA
AAGTGTGTGAAATTCATTTTCATGAAATCAACAACAAAATGGTGGAATGTAAGAAA
GCTCAGCCAAAGGAGGTGATGTCGCCAACGGGCTCAGCCCGGGGGAGGTCTCGAGT
CATGCCCTACGGAATGGACGCCTTCATGCTGGGCATCGGCATGCTGGGTTACCCAGG
TTTCCAAGCCACAACCTACGCCAGCCGGAGTTATACAGGCCTCGCCCCTGGCTACAC
CTACCAGTTCCCCGAATTCCGTGTAGAGCGGACCCCTCTCCCGAGCGCCCCAGTCCT
CCCCGAGCTTACAGCCATTCCTCTCACTGCCTACGGACCAATGGCGGCGGCAGCGGC
GGCAGCGGCTGTGGTTCGAGGGACAGGCTCTCACCCCTGGACGATGGCTCCCCCTCC
AGGTTCGACTCCCAGCCGCACAGGGGGCTTCCTGGGGACCACCAGCCCCGGCCCCA
TGGCCGAGCTCTACGGGGCGGCCAACCAGGACTCGGGGGTCAGCAGTTACATCAGC
GCCGCCAGCCCTGCCCCCAGCACCGGCTTCGGCCACAGTCTTGGGGGCCCTTTGATT
GCCACAGCCTTCACCAATGGGTACCACGCGATCGC

LgBiT-Msi1 cloning. A synthetic human Msi1 gene fragment was purchased from Twist Bioscience and inserted at the N-terminal position[9] using standard cloning techniques with EcoRV and XbaI restriction enzymes.

Gene Fragment:

5’GATATCTTATGGAGACTGACGCGCCCCAGCCCGGCCTCGCCTCCCCGGACTCGCCG
CACGACCCCTGCAAGATGTTCATCGGGGGACTCAGTTGGCAGACTACGCAGGAAGG
GCTGCGCGAATACTTCGGCCAGTTCGGGGAGGTGAAGGAGTGTCTGGTGATGCGGG
ACCCCCTGACCAAGAGATCCAGGGGTTTCGGCTTCGTCACTTTCATGGACCAGGCGG
GGGTGGATAAAGTGCTGGCGCAATCGCGGCACGAGCTCGACTCCAAAACAATTGAC
CCTAAGGTGGCCTTCCCTCGGCGAGCACAGCCCAAGATGGTGACTCGAACGAAGAA
GATCTTTGTGGGGGGGCTGTCGGTGAACACCACGGTGGAGGACGTGAAGCAATATT
TTGAGCAGTTTGGGAAGGTGGACGACGCCATGCTGATGTTTGACAAAACCACCAAC
CGGCACCGAGGGTTCGGGTTTGTCACGTTTGAGAGTGAGGACATCGTGGAGAAAGT
GTGTGAAATTCATTTTCATGAAATCAACAACAAAATGGTGGAATGTAAGAAAGCTC
AGCCAAAGGAGGTGATGTCGCCAACGGGCTCAGCCCGGGGGAGGTCTCGAGTCATG
CCCTACGGAATGGACGCCTTCATGCTGGGCATCGGCATGCTGGGTTACCCAGGTTTC
CAAGCCACAACCTACGCCAGCCGGAGTTATACAGGCCTCGCCCCTGGCTACACCTAC
CAGTTCCCCGAATTCCGTGTAGAGCGGACCCCTCTCCCGAGCGCCCCAGTCCTCCCC
GAGCTTACAGCCATTCCTCTCACTGCCTACGGACCAATGGCGGCGGCAGCGGCGGC
AGCGGCTGTGGTTCGAGGGACAGGCTCTCACCCCTGGACGATGGCTCCCCCTCCAGG
TTCGACTCCCAGCCGCACAGGGGGCTTCCTGGGGACCACCAGCCCCGGCCCCATGG
CCGAGCTCTACGGGGCGGCCAACCAGGACTCGGGGGTCAGCAGTTACATCAGCGCC
GCCAGCCCTGCCCCCAGCACCGGCTTCGGCCACAGTCTTGGGGGCCCTTTGATTGCC
ACAGCCTTCACCAATGGGTACCACTGATCTAGA

Msi2-LgBiT cloning. Musashi 2 (Msi2) was amplified from a purchased plasmid, Msi2 variant 1 in pFN21A (Promega), and inserted into pcDNA3 containing LgBiT inserted at the C-terminal position[9] using standard cloning techniques with KpnI and AsiSI restriction enzymes. Primers insert a Kozak sequence on the N-terminus.

Primers:

5’ GTACGGTACCGCCACCATGGAGGCAAATGGGAGCCAAG
5’ GTCACGGCGATCGCATGGTATCCATTTGTAAAGGCC

LgBiT-Msi2 cloning. Human Msi 2 was amplified from pcDNA3 + Msi2-Lg and ligated into a pcDNA3 construct with LgBiT inserted at the N-terminal position using standard cloning techniques with XhoI and EcoRI restriction enzymes.

Primers:

5’ TCTCCTCGAGATGGAGGCAAATGGGAGCC
5’ CAGTGAATTCTCAATGGTATCCATTTGTAAAGGCCG

C. Bioconjugation Methods

Protocol:

Pre-miRNA probes bearing a 5-aminohexylacyrlamino uridine modification and biotin appended to the 5’ end by an 18-atom spacer (1.0 mM in 100 mM phosphate buffer, pH 8.0) was mixed with an equivalent volume of HaloTag ligand (10 mM in DMSO for O2 and O4). The reaction was then allowed to proceed at 25 °C for 1 h. pre-miRNA-Cl was precipitated by the addition of 0.11× volume of 3.0 M sodium acetate (pH 5.2) and 4 volume equivalents of cold ethanol, and pelleted at 20,000 × g for 40 min at 4 °C. The pellet was then re-suspended in 100 mM phosphate buffer (pH 8.0) at a concentration of 1.0 mM and stored at -80 °C.

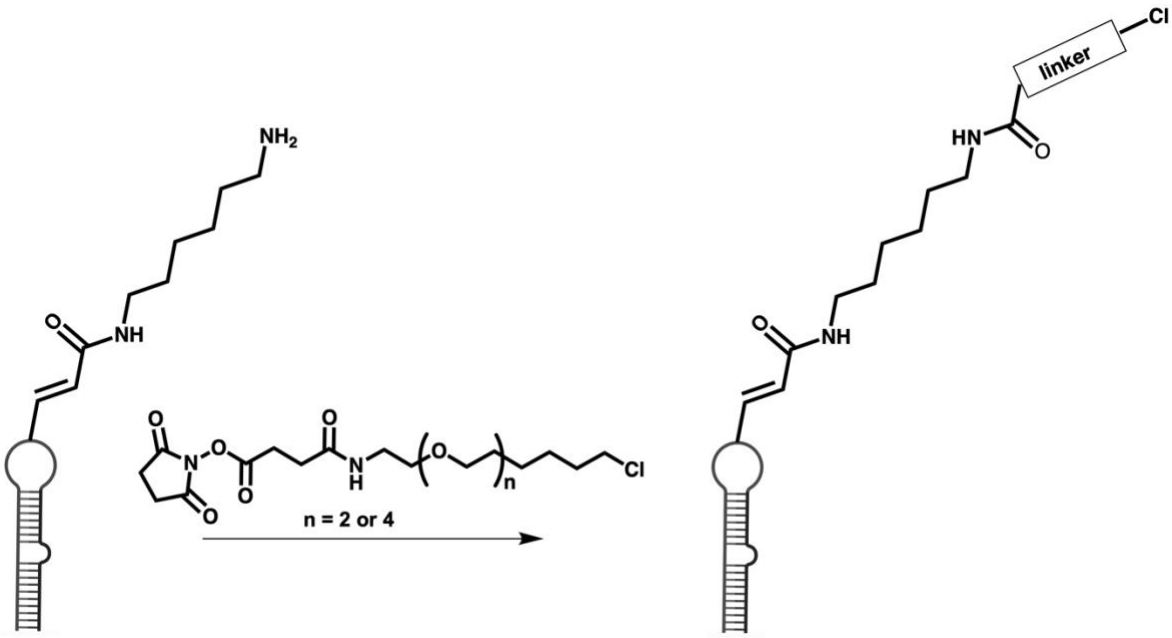


Figure 4.7. RNA labeling scheme with 5-aminohexylacrylamino uridine modifications.

D. RiPCA Protocol

Lipofectamine™ RNAiMAX protocol:

Flp-In-293 cells stably expressing a SmBiT-HT protein were reverse transfected using Lipofectamine™ RNAiMAX Transfection Reagent. Cells were passaged approximately 10 times, and no more than 15 times, before returning to low passage stocks. To test "n" number of conditions, Solution A was prepared by combining 50 × (n+1) μL of room temperature Opti-MEM™ and 2.4 × (n+1) μL plasmid encoding selected RBP-LgBiT fusion. Solution B was prepared by adding pre-miRNA-Cl and plasmid (final concentrations 0.3 μM and 0.195 ng/μL, respectively) to 50 μL Opti-MEM™ for each separate condition to be tested. Solution B was mixed with 50 μL of Solution A to yield Solution A+B, which was incubated for at least 15 min at room temperature while cells were harvested. Cells were harvested as and counted as described above. Harvested cells were used to prepare Solution C, which was composed of 400 μL of 100,000

cells/mL. Solution C was mixed with 50 µL of Solution A+B and plated 100 µL per well, four wells per condition, in a white-bottom, tissue culture-treated 96-well plate (Corning cat #3917). The plate was incubated in a humidified incubator (37 °C and 5% CO_2) for 24 h. After incubation, the media was removed and replaced with 100 µL room temperature Opti-MEM™ and treated with 25 µL NanoGlo Live Cell Reagent diluted 1:20 according to the manufacturer's recommendation. All chemiluminescence data was collected immediately on a BioTek Cytation3 plate reader.

Representative calculations based on an assay for n = 5 conditions:

<u>Solution A: Prepared for n+1= 6</u>

6 x 50 µL → 300 µL Opti-MEM™

6 x 2.4 µL → 14.4 µL Lipofectamine™ RNAiMAX

<u>Solution B:</u>

50 µL Opti-MEM™

2.5 µL 3.9 ng/µL RBP-LgBiT plasmid

0.3 µL 50 µM pre-miRNA-Cl

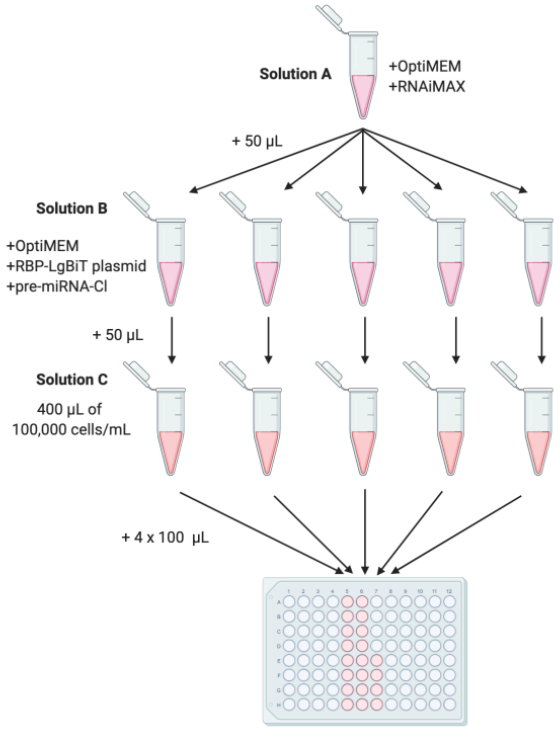


Figure 4.8. Lipofectamine™ RNAiMAX transfection workflow.

TransIT-X2® Protocol:

Flp-In-293 cells stably expressing a SmBiT-HT protein were reverse transfected using TransIT-X2® Reagent. Cells were passaged approximately 10 times, and no more than 15 times, before returning to low passage stocks. Solution B for each condition was prepared by adding in order DNA (volumes provided in Table S1), 0.45 μL of 25 μM RNA probe, and 1.126 μL TransIT-X2® to 37.5 μL room temperature Opti-MEM™. Solution B was mixed by briefly vortexing and was briefly centrifuged prior to ~15 min incubation at room temperature while cells were harvested. Cells were harvested as and counted as described above. Harvested cells were used to prepare Solution C (300 μL × (n+1) of 100,000 cells/mL) and 300 μL of Solution C was added to Solution B. Solution B+C was mixed by pipetting up and down before plating 100 μL per well, 3 wells per condition, in a white-bottom, tissue culture-treated 96-well plate (Corning cat #3917). The plate was incubated in a humidified incubator (37 °C and 5% CO_2) for 24 h. After incubation, the media was removed and replaced with 100 μL room temperature Opti-MEM™ and treated with 25 μL NanoGlo Live Cell Reagent diluted 1:20 according to the manufacturer's recommendation. All chemiluminescence data was collected immediately on a BioTek Cytation3 plate reader.

Table 4.1. Volumes, concentrations, and amounts of DNA used with each of the RBPs in RiPCA 2.0.

RBP	Volume of DNA	Concentration of DNA	Amount of DNA per well
Lin28A/B	0.47 μL	3.9 ng/μL	~0.5 ng
hnRNP A1	1.876 μL	3.9 ng/μL	~2 ng
Msi1/2	1.726 μL	3.9 ng/μL	~2 ng

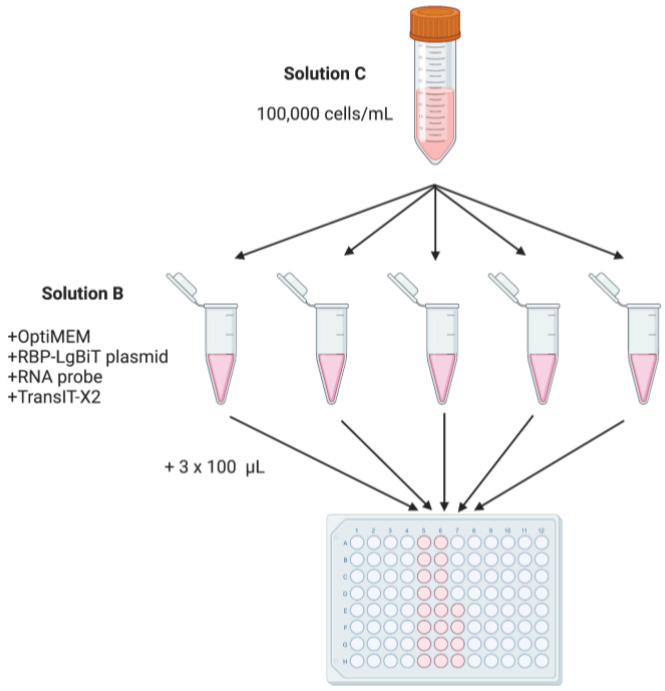


Figure 4.9. TransIT-X2® transfection workflow.

Cell Titer Glo Protocol:

Cells were transfected following the Lipofectamine™ RNAiMAX or TransIT-X2® protocol and incubated in a white-bottom, tissue culture-treated 96-well plate (Corning cat #3917) for 24 h in a humidified incubator (37 °C and 5% CO_2). At 24 h, media was removed and replaced with 70 µL Opti-MEM™ and 70 µL Cell Titer Glo reagent reconstituted according to the manufacturer's protocol. The plate was incubated for 30 min then chemiluminescence signal was collected on a BioTek Cytation3 plate reader.

SmHT EMSA Protocol:

Pre-miRNA probe (500 nM) and purified SmHT[10] (5 µM) were incubated in 10 µL PB8 for 1 h at room temperature. Reaction was quenched with 10 µL 2X RNA loading dye (95% formamide, 0.02% SDS, 0.02% bromophenol blue, 0.01% xylene cyanol, 1 mM EDTA in H_2O) and ran on

TBE-Urea gels (10% for 45 min at 200V or 15% for 60 min at 200V). Gels were imaged with ChemiDoc™ Imaging System (Biorad).

4.6 References

1. Treiber, T.; Treiber, N.; Plessmann, U.; Harlander, S.; Daiss, J. L.; Eichner, N.; Lehmann, G.; Schall, K.; Urlaub, H.; Meister, G., A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Molecular Cell* **2017,** *66* (2), 270-284.

2. Kooshapur, H.; Choudhury, N. R.; Simon, B.; Muhlbauer, M.; Jussupow, A.; Fernandez, N.; Jones, A. N.; Dallmann, A.; Gabel, F.; Camilloni, C.; Michlewski, G.; Caceres, J. F.; Sattler, M., Structural basis for terminal loop recognition and stimulation of pri-miRNA-18a processing by hnRNP A1. *Nat Commun* **2018,** *9* (1), 2479.

3. Michlewski, G.; Caceres, J. F., Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nature Structural & Molecular Biology* **2010,** *17* (8), 1011-1128.

4. Michlewski, G.; Guil, S.; Cáceres, J. F., Stimulation of pri-miR-18a Processing by hnRNP A1. In *Regulation of microRNAs*, Großhans, H., Ed. Springer US: New York, NY, 2010; pp 28-35.

5. Kawahara, H.; Okada, Y.; Imai, T.; Iwanami, A.; Mischel, P. S.; Okano, H., Musashi1 cooperates in abnormal cell lineage protein 28 (Lin28)-mediated let-7 family microRNA biogenesis in early neural differentiation. *J Biol Chem* **2011,** *286* (18), 16121-30.

6. TransIT-X2® Dynamic Delivery System: A dynamic transfection system for the delivery of plasmid DNA, siRNA/miRNA and CRISPR/Cas9 components. https://www.mirusbio.com/products/transfection/transit-x2-dynamic-delivery-system.

7. Siomi, H.; Dreyfuss, G., A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol* **1995,** *129* (3), 551-60.

8. Li, N.; Yousefi, M.; Nakauka-Ddamba, A.; Li, F.; Vandivier, L.; Parada, K.; Woo, D. H.; Wang, S.; Naqvi, A. S.; Rao, S.; Tobias, J.; Cedeno, R. J.; Minuesa, G.; Katz, Y.; Barlowe, T. S.; Valvezan, A.; Shankar, S.; Deering, R. P.; Klein, P. S.; Jensen, S. T.; Kharas, M. G.; Gregory, B. D.; Yu, Z. Q.; Lengner, C. J., The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. *Cell Reports* **2015,** *13* (11), 2440-2455.

9. Rosenblum, S. L.; Lorenz, D. A.; Garner, A. L., A live-cell assay for the detection of pre-microRNA–protein interactions. *RSC Chemical Biology* **2021,** (2), 241-247.

10. Sherman, E. J.; Lorenz, D. A.; Garner, A. L., Click Chemistry-Mediated Complementation Assay for RNA-Protein Interactions. *ACS Comb Sci* **2019,** *21* (7), 522-527.

CHAPTER 5

Identification of Small Molecule Inhibitors of let-7/Lin28 with RiPCA

In Chapter 1, the biological significance of targeting the pre-miRNA/RBP interaction between the let-7 family of miRNA and Lin28 was discussed. Chapter 2 highlighted several *in vitro* assay platforms used to identify small molecule inhibitors of RPIs that have been employed to discover inhibitors of let-7/Lin28, the results of which will be further discussed in this chapter. Also mentioned in Chapter 2 were cell-based assays that offer many benefits to high-throughput screens (HTS), though, to-date, none have been used to screen for inhibitors of pre-miRNA/RBP interactions. RiPCA 2.0 eliminates the limitations of these cell-based assays and, therefore, presents a promising platform to screen for inhibitors of let-7/Lin28. RiPCA 2.0 was adapted for HTS using liquid handling at the University of Michigan Center for Chemical Genomics (CCG) and utilized to screen a library of ~18,000 small molecules curated and provided by Merck. This chapter reports assay optimization, implementation, and follow up, including confirmation assays and dose-response and toxicity analysis, that yielded a set of seven compounds for further characterization.

5.1 Previously Reported Lin28 Inhibitors

As mentioned in Chapter 2, several assay platforms have been adapted to screen for inhibitors of the let-7/Lin28 RPI. The first to do so was Roos *et al.* in 2016 using a FRET-based assay in which they screened a library of 16,000 small drug-like molecules. Of the initial 203 hits

that were re-evaluated in triplicate and corrected for compound autofluorescence, 14 compounds were followed through to and further triaged by cellular evaluation in assays such as a luciferase gene reporter assay, RT-qPCR, and colony formation assays. These efforts highlighted compound **1632**, which was found to bind to Lin28 and shown to inhibit stemness and induce differentiation in murine ESCs and inhibit proliferation in cancer cell lines (Fig. 5.1). In Huh7 cells, a dose-dependent increase in let-7a, let-7d, and let-7f levels was observed with **1632**. However, this compound was originally reported as an anxiolytic with high affinity for benzodiazepine receptor sites in the brain and demonstrated activity against the bromodomains of BRD4 and CREBBP with activities comparable to that observed with Lin28A (K_ds of 7 and 25 μM for BRD4 and CREBBP, respectively, and Lin28A IC_{50} of 8 μM.[1] This is notable in light of the finding that Lin28A expression is reduced in the presence of the bromodomain inhibitor JQ1 reported by Sin-Chan *et al.*[2]

The same year, Lim *et al.* reported another FRET-based screen of a 4,500-compound library of small molecules against let-7/Lin28. Primary screening hits were confirmed by EMSA, which yielded a single hit, **1**, that was followed up in additional biochemical and cellular assays, including SPR, RT-qPCR, RNA pull-down (Fig. 5.1). From these assays, it was determined that the hit molecule bound to the CSD of Lin28, inhibiting its interaction with the loop region of pre-let-7a-1 and -7g with IC_{50} values in the low μM range and freeing pre-let-7g for cleavage by Dicer *in vitro*. Cellular and Lin28-dependent activity of the compound was confirmed by RT-qPCR, which detected increased levels of let-7s, but not other miRNAs, in JAR cells treated with the compound. Finally, a modest decrease in let-7 target genes was noted by a luciferase reporter gene assay and Western blot.[3]

Lightfoot *et al.*, Wang *et* al, and Borgelt *et al*. utilized FP-based assays to identify inhibitors of let-7/Lin28. Lightfoot *et al.* screened 2,768 pharmacologically active small molecules, which yielded 64 primary hits. After re-testing of 44 filtered hits for reproducibility, 21 hits were selected and validated in biochemical assays, namely EMSA and Dicer activity assay. Based on the activity of hits in EMSA, four compounds, 4, 10, 11, and 14, were determined to be true hits with IC_{50}

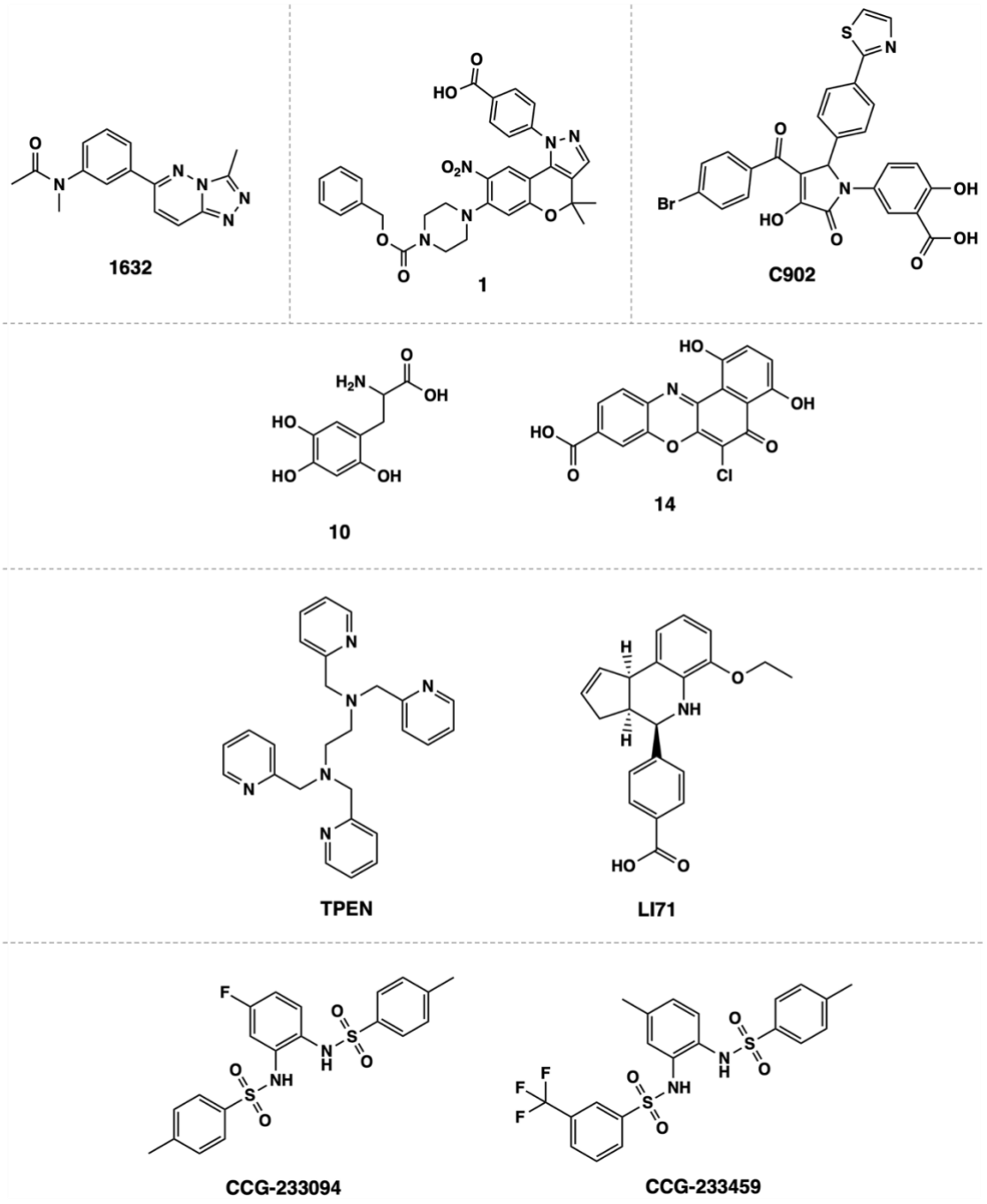


Figure 5.1. Structures of previously reported Lin28 inhibitors.

values in the low μM range. Only **10** and **14** restored Dicer processing in the presence of Lin28 *in vitro* (Fig. 5.1). However, the cellular activity of these molecules was not interrogated.[4]

Wang *et al.* developed a similar FP-based approach to screen 17 compound libraries, with a total of 101,017 small molecules, against Lin28 using a truncated pre-let-7f-1 sequence. From the primary screen, 350 molecules were selected for dose-response testing, which yielded 53 compounds with IC_{50} values ranging from 200 nM to 10 μM. Of the 53 compounds, 42 were commercially available and re-confirmed in dose-response. The resulting 27 hits were further filtered and assessed for their ability to inhibit oligouridylation of pre-let-7g. Two compounds, **TPEN** and **LI71**, were selected for further characterization of mechanism of action and assessment in cellular assays (Fig. 5.1). TPEN is a known cell permeable Zn^{2+} chelator capable of inducing apoptosis. Unsurprisingly, TPEN was found to inhibit Lin28 via binding to the ZKD. LI71, on the other hand, was determined to interact with the Lin28 CSD and shown to have activity against Lin28 in leukemia and embryonic stem cells. TPEN was not accessed for cellular activity, but LI71 was found to increase the levels of let-7 family members to varying degrees in K562 and DKO+A mESCs.[5]

In 2021, Borgelt *et* al. reported an FP screen of ~15,000 natural product-like small molecules against Lin28 and the terminal loop of pre-let-7f-1. The primary screening hits were filtered by activity in FP as well as purity by LC-MS. Six heterocyclic small molecules were selected as the initial hits, one of which, **C902**, showed dose-dependent inhibition in the micromolar range and was carried to further characterization. The inhibitory activity of C902 was validated in EMSA and was determined to stabilize the Lin28 CSD. Treatment of JAR cells with C902 yielded a ~2-fold increase in let-7a and let-7g levels as quantified by RT-qPCR. C902 is a trisubstituted pyrroline and part of a series of molecules known to stabilize the protein-protein

interaction involving 14-3-3 plant proteins, which was shown by Borgelt *et al.* to be tunable. In SAR studies, it was shown that compounds tended to favor either RPI or PPI inhibitory activity.[6]

Diverging from fluorescence-based methods, the Garner laboratory utilized cat-ELCCA to screen 127,007 compounds from several small molecule libraries against full length pre-let-7d/Lin28. The primary screen yielded 1,468 compounds that were re-tested in triplicate. Following dose-response analysis of 136 re-confirmed compounds with IC_{50} values ranging from 0.01 to 100 μM, additional selection criteria, including known promiscuous binding, reactivity, and tractability of medicinal chemistry was used to identify 20 top hits for further follow up studies. After purchasing the selected compounds, 10 retained inhibitory activity, but only two structurally related compounds, **CCG-233094** and **CCG-234459**, demonstrated dose-dependent inhibition in cat-ELCCA (IC_{50} values of 8.3 μM and 10.3 μM, respectively) and EMSA (Fig. 5.1). Using SPR, it was determined that the compounds bound to Lin28 and there was minimal disruption of Dicer activity in a gel-based assay. In subsequent SAR studies, most modifications to the structure reduced activity. Unfortunately, cellular assessment of hit compounds did not yield significant cellular activity and the hits displayed cytotoxicity.[7]

Common to the numerous efforts to develop small molecule inhibitors of let-7/Lin28 is the identification of protein-binding small molecules. To date, none of the reported compounds have demonstrated RNA-binding. Additionally, there is no specificity of inhibition as increases in the levels of multiple let-7 family members is observed upon treatment of several reported molecules[1, 3, 5-6], which indicates non-specific inhibition of Lin28-RNA binding, an undesirable property when attempting to inhibit the let-7 interaction alone. Furthermore, all have micromolar activity in biochemical assays and demonstrate reduced potency in cells. Additionally, target engagement in cells was not confirmed in all reports, and in some cases revealed significant off-target

interactions[1].Therefore, it was appealing to utilize RiPCA as a cell-based inhibitor discovery platform and screen a library of RNA-biased compounds.

5.2 Adaptation of RiPCA to HTS

In its original format, RiPCA was performed in 96-well plates. To reduce reagent consumption and enable HTS, RIPCA was miniaturized to 384-well format. Based on the strong performance of pre-let-7d-36-4Cl in RiPCA 2.0, this probe was selected as the substrate in the screen. Several points of optimization and adaptation to liquid handling at CCG were required for miniaturization. First, the optimal number of cells per well was determined. Of the four conditions tested, 3,000 cells per well produced the largest S/B (Fig. 5.2—by hand). The next point of optimization involved the media removal step immediately prior to the addition of Opti-MEM™ and NanoLuc substrate. It is essential to remove as much of the cell culture media as consistently as possible since leftover DMEM causes an increase in signal detected. Therefore, media removal by plate washer and the Biomek liquid handler was assessed. The three conditions were tested in clear-bottom 384-well plates to enable tandem visualization of the cells and RiPCA assay reading after media removal and replacement. Interestingly, removal of media by hand smacking the plate

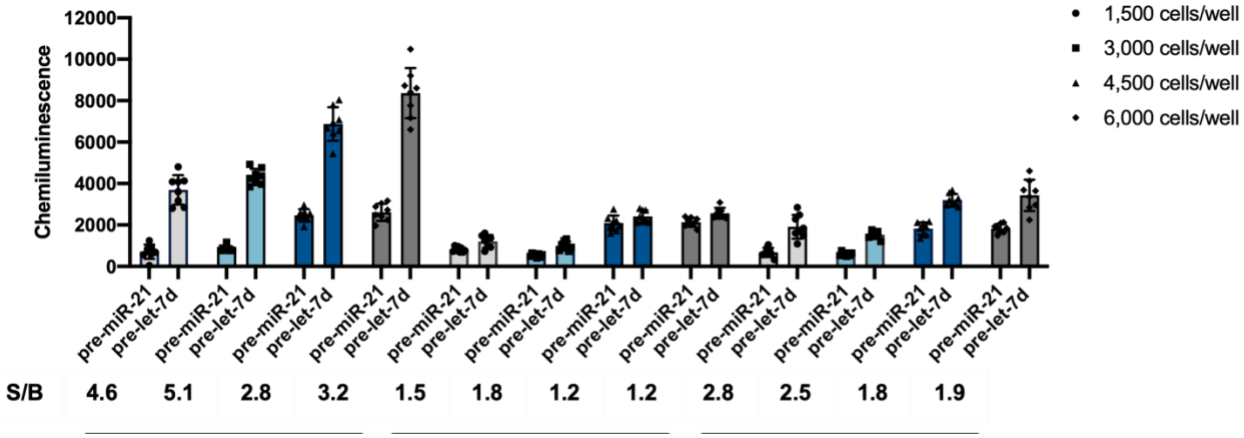
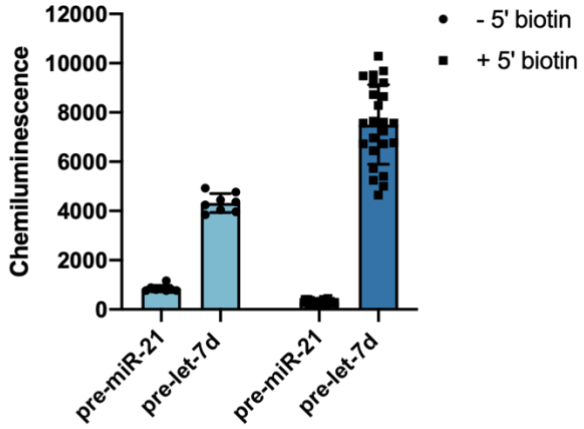


Figure 5.2. Optimization of RiPCA in 384-well format. Comparison of methods of media removal, by hand, with a plate washer, or with the Biomek liquid handler, prior to addition of signal detection reagents.

on a spill pad afforded the greatest retention of cells in each well with corresponding S/B, whereas very few cells remained in the wells when media was removed mechanically (Fig. 5.2).



S/B	5.1	25.7
Z'	0.56	0.31

Figure 5.3. Testing effect of biotin in pre-miRNA probe. Comparison of signal and S/B generated with pre-miRNA probes with and without a 5' biotin modification.

In these initial proof-of-concept experiments, a pre-let-7d probe lacking a 5' biotin modification was used. Given that the S/B observed was significantly lower than seen in previous RiPCA 2.0 experiments, the corresponding biotinylated probe was tested in 384-well format. The 5' biotinylated pre-let-7d probe produced ~5-fold greater S/B (Fig. 5.3). It is possible that the 5' biotin protects the hairpin from degradation, or the physiochemical properties of biotin enhance transfection. It was previously confirmed that placing the HaloTag ligand on the 5' end of pre-let-7d does not disrupt the Lin28 binding it RiPCA (Fig. 3.5) and 5' labeled pre-let-7d was utilized in the Lin28 cat-ELCCA screen[7]; therefore, biotinylated pre-let-7d-4Cl was used as the optimized assay reagent.

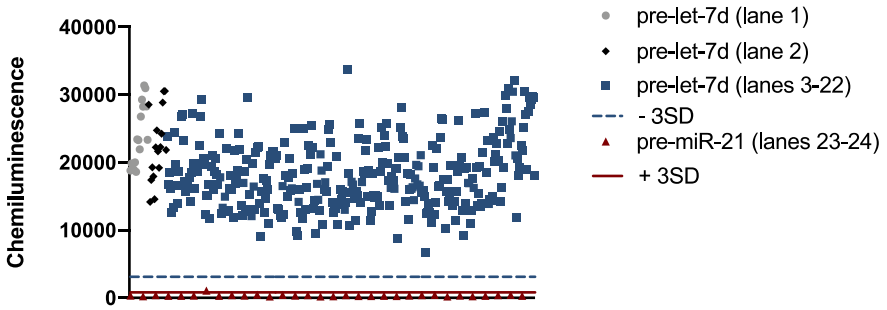


Figure 5.5. Full 384-well plate test. Distribution of signal generated by pre-let-7d (blue squares) and pre-miR-21 (red triangles) wells in a 384-well plate. Blue dotted line indicates 3 standard deviations below the mean of pre-let-7d (negative control) signal and the red solid line indicates 3 standard deviations above the pre-miR-21 (positive control) signal.

Next, the ability of RiPCA to be scaled up to plate an entire 384-well plate was assessed. The transfection protocol was proportionally scaled up to 32 times the normal volume and plated in lanes 1-22 of a 384-well plate using a Multidrop Combi Reagent Dispenser to aliquot 30 μL per well. A pre-miR-21 batch transfection, intended to serve as the positive control in the screen, was plated in the remaining two lanes. Lanes 1 and 2 were designated as negative control lanes to be treated with DMSO only. The average S/B for the entire plate was 51.7 with a Z' of 0.16 (Fig. 5.4). The Z' is a metric used to assess the quality of an assay. Practically, it is a measure of the assay to differentiate positive and negative results utilizing the average and standard deviation of positive and negative controls. It is standard to require an assay's Z' to be > 0.5 to proceed as an HTS platform. However, Bar and Zweifach dispute this and suggest that this is too stringent of a requirement for phenotypic and cell-based assays, which by nature produce more variable results.[8]

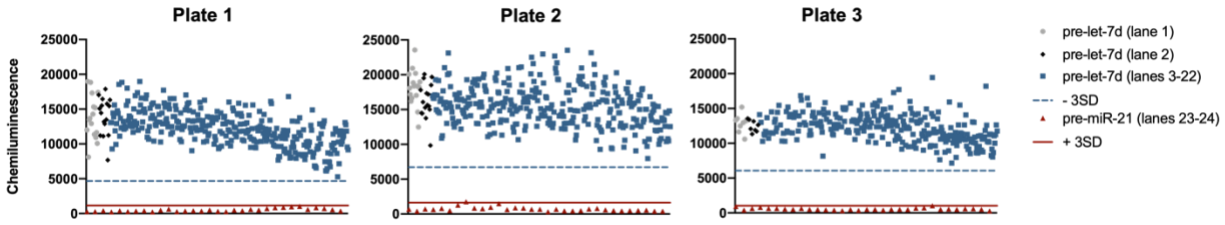


Figure 5.4. Scaled up transfection test. Distribution of signal generated by large batch transfected pre-let-7d (blue squares) and pre-miR-21 (red triangles) wells in three 384-well plates. Blue dotted line indicates 3 standard deviations below the mean of pre-let-7d (negative control) signal and the red solid line indicates 3 standard deviations above the pre-miR-21 (positive control) signal.

To increase throughput, the batch transfection was scaled up even further to 77 times original reagent quantities, which was sufficient for plating three full 384-well plates. While the average S/B decreased to ~23, the Z' increased to an average of 0.36 across the three plates (Fig. 5.5). Encouraged by these results and to practice the HTS workflow (Fig. 5.6), a pilot screen was performed using a plate from the LOPAC library. At 16 h post-transfection, 10 µL of 40 µM compound in DMEM was added to each well using a Biomek liquid handler from an intermediate compound plate prepared by the addition of 60 nL DMSO or compound by the ECHO liquid

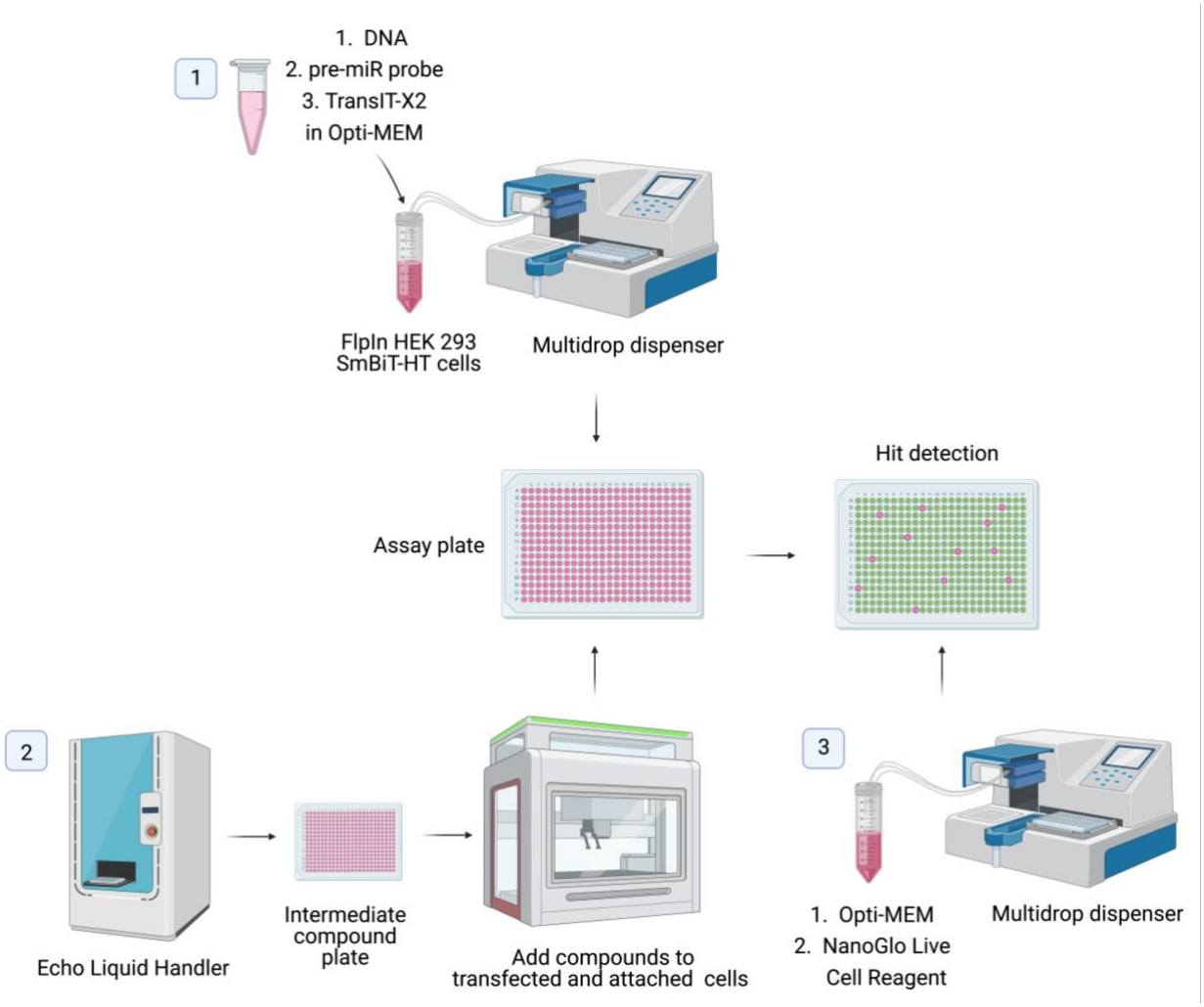


Figure 5.6. RiPCA HTS workflow. (1) Batch transfected FlpIn-293 SmBiT-HT expressing cells are plated in a 384-well assay plate using a Multidrop dispenser. **(2)** An intermediate compound plate is prepared using an ECHO liquid handler and volume is transferred to the assay plate using a Biomek liquid handler. **(3)** Chemiluminescence is measured after media removal and replacement with Opti-MEM™ and NanoGlo Live Cell Reagent.

handler and 15 μL DMEM. The assay plate was returned to the incubator for 8 h and chemiluminescence signal was measured 30 min after the addition of the NanoLuc substrate. Of the 320 compounds tested, there were four that reduced signal greater than 3 standard deviations below the mean signal for the negative control lanes (lanes 1-2) for a preliminary hit rate of 1.25% (Fig. 5.7).

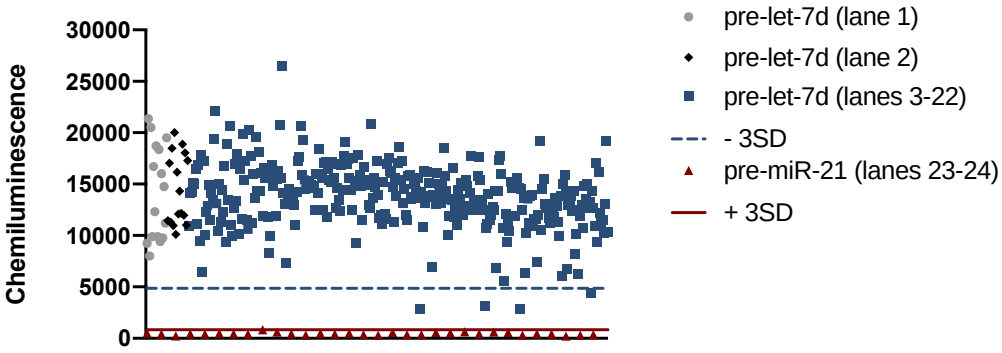


Figure 5.7. Pilot screen. Screened 320 compounds from LOPAC library to test HTS protocol. Blue dotted line indicates 3 standard deviations below the mean of pre-let-7d (negative control) signal and the red solid line indicates 3 standard deviations above the pre-miR-21 (positive control) signal.

5.3 Lin28A RiPCA High-throughput Screen

In an effort to narrow the chemical space screened to structures that are likely to bind RNA, the compounds screened belonged to an RNA-biased small molecule library curated by Merck and included compounds identified in a virtual screen conducted at Merck against published structures of Lin28 bound to the terminal loop of pre-let-7s. Using the optimized HTS workflow, the collection of 17,797 small molecules was screened at 10 μM. Due to consistent, excessive variability in the first column of each plate, the data from lane 1 was eliminated from data analysis and lane 2 was utilized as the negative control. The average S/B and Z' of the campaign calculated from the negative control in lane 2 and positive controls in lanes 23 and 24 were 14.7 and 0.37, respectively. Compounds that demonstrated > 55% inhibition were selected as primary hits (Fig.

5.8B). The 992 selected compounds (5.6% primary hit rate) were then counterscreened in triplicate at 10 μM for cytotoxicity with Cell Titer Glo. Compounds that exhibited toxicity of >20% in at least two of three were eliminated, yielding a hit list of 840 compounds that were re-confirmed in RiPCA in triplicate (Fig. 5.8C).

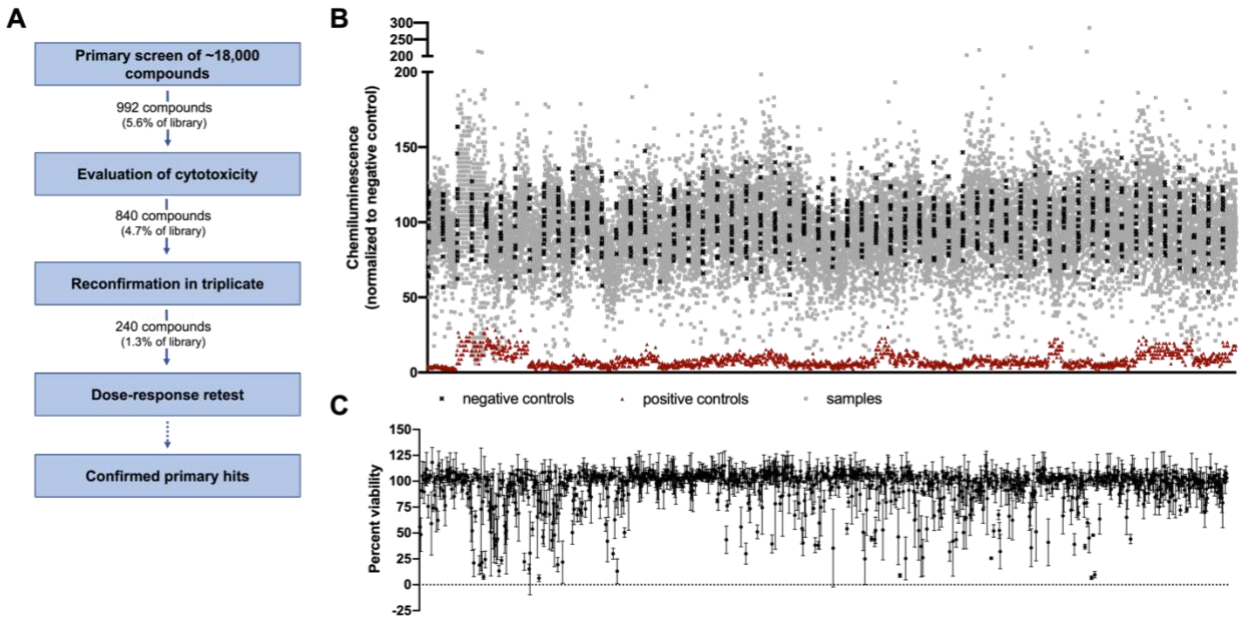


Figure 5.8. Lin28 RiPCA HTS Screen. (A) Screening funnel for Lin28 RiPCA HTS. **(B)** Primary screen of 17,797 compounds in RiPCA. Compounds were tested at 10 μM for 8 hr. **(C)** Viability counterscreen of 840 primary hits. Data displayed as mean ± SD.

Of the 840 compounds tested, 240 (28.6%) demonstrated inhibitory activity of >55% in 2 of the 3 replicates. These compounds were then tested in dose-response with a concentration curve ranging from 0.391 to 50 μM. Roughly one-third of tested molecules displayed concentration-dependent inhibition, which was defined as pAC_{50} value > 5. To build in SAR, 239 additional compounds with structural similarity to the 78 hits were included in subsequent assays. The selected 317 compounds were again tested in triplicate at 10 μM for cytotoxicity in Cell Titer Glo. The stringency of the selection criteria was decreased at this step and compounds exhibiting average toxicity >40% were eliminated, yielding 264 compounds, all of which were then tested in dose-response in RiPCA as well as cat-ELCCA, used as a counterscreen for on-target let-7/Lin28

inhibition. In cat-ELCCA, only 8.7% of tested compounds exhibited concentration-dependent inhibition let-7/Lin28. All 23 compounds for which a value was determined had pAC_{50} values between 3.07 and 5.52. Interestingly, for all but one compound the pAC_{50} value was greater in RiPCA than cat-ELCCA. From these compounds, a representative subset of seven compounds from three chemical series were chosen for further characterization (Fig. 5.9A-C).

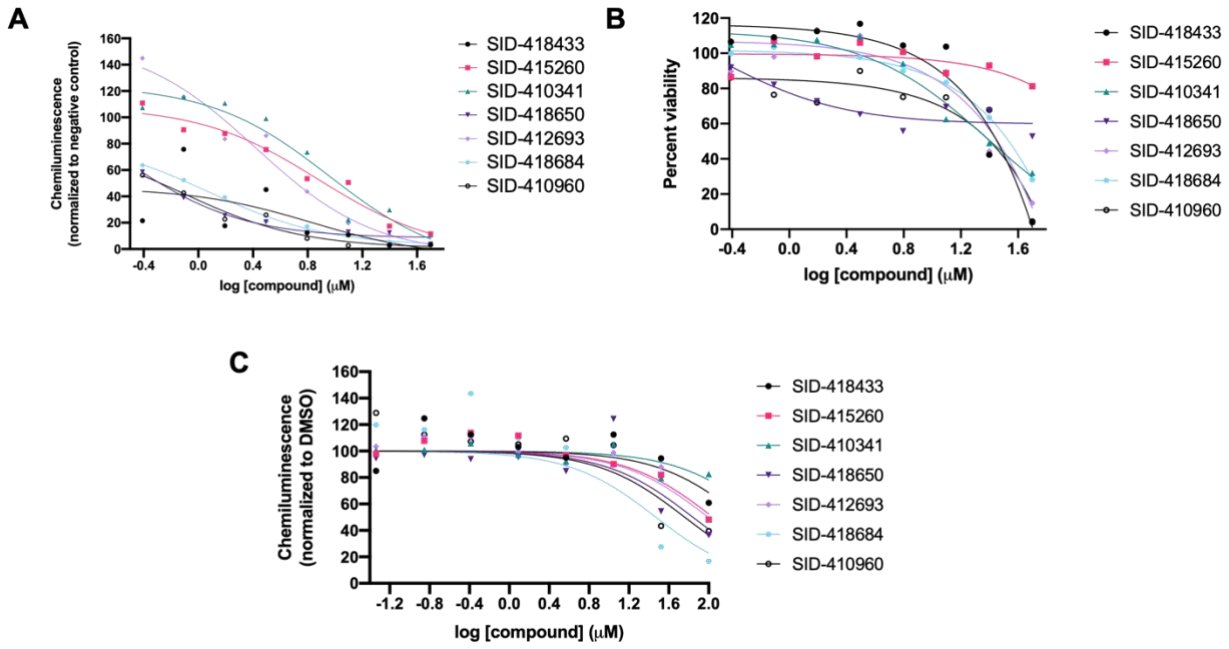


Figure 5.9. Evaluation of top hits in RiPCA, Cell Titer Glo, and cat-ELCCA. Top seven hits were evaluated in dose-response in **(A)** RiPCA, **(B)** Cell Titer Glo, and **(C)** cat-ELCCA.

5.4 Biochemical and Cellular Characterization of Top Hits

To determine whether compounds 1-7 could relieve Lin28A-mediated inhibition of let-7, the compounds were evaluated in a Dicer digest assay as well as with RT-qPCR. Pre-let-7d-4-Cl was pre-incubated with 50 μM compound for 15 min then Dicer was allowed to digest the RNA for 2 h at 37 °C. Dicer products were analyzed on a TBE-Urea gel. Figure 5.10B shows that compound 3 displays no inhibitory activity against Dicer, whereas compounds 1, 2, and 5 display partial inhibition and 4, 6, and 7 show full inhibition of Dicer processing. To assess cellular

activity, the compounds were evaluated in T47d cells, a cell line cultured from a breast cancer patient, that express moderate levels of Lin28A. Let-7 levels were quantified after 48 h treatment with 5 μM via RT-qPCR. It is important to note the difference in concentration of compound tested is due to high toxicity of compounds in the cell line tested at higher concentrations. At this concentration, there was slight increase in let-7a and let-7d levels upon treatment with compound 2 (Fig. 5.10C). Compound 6 was only evaluated in one replicate due to insufficient RNA harvested due to toxicity in the second replicate (Fig. 5.10C).

A

Compound #	SID-#
1	SID-418433
2	SID-415260
3	SID-410341
4	SID-418650
5	SID-412693
6	SID-418684
7	SID-410960

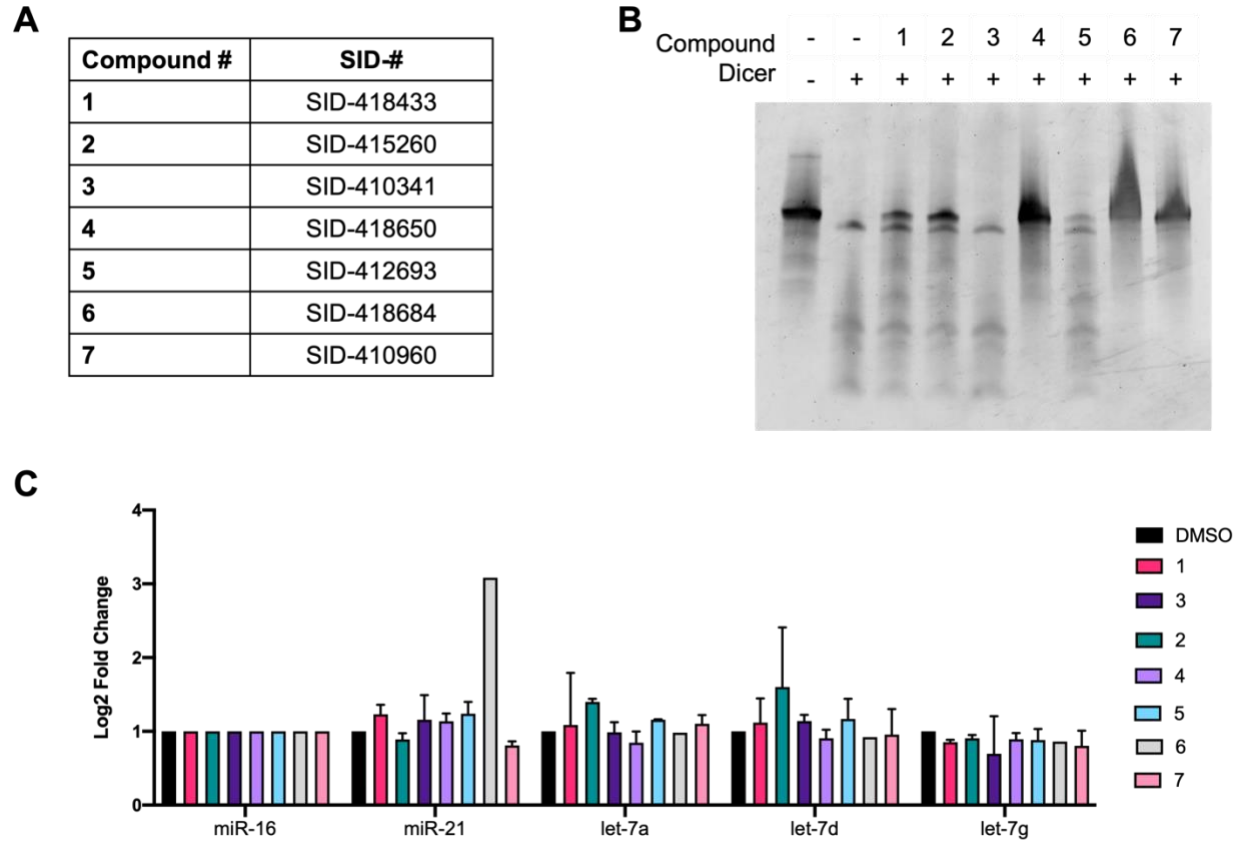


Figure 5.10. Evaluating top hits. (A) List of compound numbers and their associated SID-#. **(B)** Inhibition of Dicer processing of pre-let-7d by compounds 1-7 (50 μM). **(C)** Let-7 levels after 48 h treatment with 5 μM compound quantified by RT-qPCR (n = 2).

Further analysis is required to fully evaluate and characterize compounds 1-7. For example, it would be desirable to determine the mode and site of compound binding as well as assess the

compounds at higher concentrations and in other cell lines, such as JAR and Huh7, that have been used in other published studies.[1, 3]

5.5 Conclusions

The work outlined in this chapter presents the successful adaptation and implementation of RiPCA HTS for the discovery of cellular inhibitors of let-7/Lin28. While additional characterization is required, the scaffolds identified by RiPCA serve as a promising starting point for further probe or inhibitor development. Together with the data presented in Chapter 4, this demonstrates the utility of RiPCA as a platform for the discovery of inhibitors for other pre-miRNA-protein interactions. Future uses and further development of RiPCA will be discussed in Chapter 6.

5.6 Methods

Materials:

Chemically synthesized pre-microRNAs (deprotected, desalted and HPLC purified), containing 5-aminohexylacrylamino uridine modifications and biotin attached to the 5'-end of the sequence by an 18-atom spacer, were purchased from Dharmacon and used as received for the labeling reaction. HaloTag Succinimidyl Ester (O4) Ligand was purchased from Promega and used as received (cat #P6751). Note that the HaloTag Succinimidyl Ester Ligands should be dissolved and immediately portioned into single use aliquots stored at -80 °C to avoid degradation. Flp-In™-293 cells and associated vectors were purchased from ThermoFisher Scientific (Invitrogen cat #75007 and #601001, respectively). The Nano-Glo Live Cell Assay System was purchased from Promega and used as received (cat #N2012). Lipofectamine™ RNAiMAX (Invitrogen cat #13778100), TransIT-X2 (Mirus cat #6000), Ni-NTA Agarose resin (Qiagen cat #30210), Streptavidin Coated

High Capacity 384-well plates (Thermo cat #15505), SuperSignal™ West Pico PLUS (Thermo cat #34578), and Horseradish peroxidase (HRP) (Pierce cat #PI31491) were used as received. 10X PBS (Invitrogen cat #AM9625) was diluted to 1X in Milli-Q H_2O.

General Cell Culture Methods:

Flp-In™-293 cells stably expressing either SmBiT-HaloTag were grown in DMEM (Corning cat #10-017-CV) supplemented with 10% FBS (Atlanta Biologicals S11550), L-glutamine (Gibco cat #25030081), hygromycin B (100 μg/mL) (Gibco cat #10687010), and Penicillin-Streptomycin (100 U/mL) (Gibco cat #15140122) at 37 °C with 5% CO_2 in a humidified incubator, passaged at least once before use for an experiment. T47d cells were grown in RMPI (Gibco cat #11875-093) supplemented with 10% FBS (Atlanta Biologicals S11550), 1 mM sodium pyruvate (Sigma cat #8636-100ML), and 7 mg/mL human insulin (Santa Cruz Biotechnology cat #360248) at 37 °C with 5% CO_2 in a humidified incubator, passaged at least once before use for an experiment. Cells were passaged using Trypsin-EDTA (0.25%) (Gibco cat #25300054) approximately 10 times, and no more than 15 times, before returning to low passage stocks. To count cells, cells were harvested and 10μL of the cell suspension was mixed with 10μL Trypan Blue (Gibco cat #15250061) ([final] = 0.2% trypan blue) and counted using a hemocytometer.

General assay and data analysis methods:

Chemiluminescence data was collected on a PerkinElmer EnVision or BMG Labtech PHERAStar plate reader. All data was analyzed using GraphPad Prism version 9.0.0 for Mac OS X (GraphPad Software, www.graphpad.com).

Bioconjugation Methods:

Pre-miRNA probes

Pre-miRNA probes (1.0 mM in PB8; 100 mM phosphate buffer, pH 8.0) were mixed with an equivalent volume of HaloTag ligand (10 mM in DMSO) for RiPCA or trans-cyclooctene (TCO) (10 mM in DMSO). The reaction was then allowed to proceed at 25 °C for 1 h. Pre-miRNA-Cl was precipitated by the addition of 0.11× volume of 3.0 M sodium acetate (pH 5.2) and 4 volume equivalents of cold ethanol, and pelleted at 20,000 × g for 40 min at 4 °C. The pellet was then re-suspended in 100 mM phosphate buffer (pH 8.0) at a concentration of 1.0 mM and stored at -80 °C.

mLin28A-HaloTag-Biotin Expression, Purification and Biotinylation

E. coli BL21 cells transformed with pFN29k + mLin28A-HaloTag were grown in 1 L LB at 37 °C to an OD_{600} of 0.94 and induced with 100 μM IPTG and grown overnight at 18 °C for 18 hr. Cells were pelleted for 30 min at 4000 rpm, resuspended in 35 mL Lysis Buffer (20 mM imidazole pH 8, 10 mM phosphate, 2.7 mM KCl, 137 mM NaCl, fresh 0.1% PMSF and 1 mM DTT), lysed by sonication, and re-centrifuged at 4000 rpm for 30 min. The supernatant was added to 5 mL of NiNTA resin equilibrated with Lysis Buffer and incubated while rotating at 4 °C for 20 min. Resin was washed with 45 mL Lysis Buffer, 45 mL 25:75 Lysis:Wash Buffer (Wash Buffer; 10 mM Tris pH 8, 50 mM imidazole, 500 mM NaCl, fresh 0.1% PMSF and 1 mM DTT). mLin28A-HaloTag was eluted from the resin in 2 × 15 mL Elution Buffer (10 mM Tris pH 8, 500 mM imidazole, 500 mM NaCl, fresh 0.1% PMSF and 1 mM DTT). Elutions were dialyzed overnight into Storage Buffer (20 mM Tris pH 8, 100 mM KCl, 0.2 mM EDTA, 10% glycerol). Dialyzed protein was incubated with 20 equivalents of biotin-PEG7-Cl overnight then buffer exchanged into Storage Buffer, quantified by BCA using BSA standards, and stored at –80 °C.

mTet-HRP

HRP (13.5 mg/mL in PBS; 100 mM phosphate, 150 mM NaCl, pH 7.0) was incubated with 7.1 mM mTet-NHS dissolved in DMSO in a total volume of 200 μL and gently shaken at room temperature for 3 hr. mTet-HRP was concentrated in PBS to remove DMSO and unreacted mTet-NHS, quantified, and stored at 4 °C.[9]

RiPCA Screening Protocol

384-well Protocol

A batch transfection solution was prepared by mixing Opti-MEM™, DNA, and RNA then adding TransIT-X2. For each 50 μL Opti-MEM™, 0.625 μL of 3.9 ng/μL pcDNA3 + mLin28A-LgBiT, 0.3 μL of 50 μM pre-miRNA-Cl probe, and 2.4 μL TransIT-X2 were added. The transfection solution was vortexed and briefly centrifuged then incubated at room temperature for ~15 min while cells were harvested. Flp-In-293 SmHT cells were harvested following standard cell culture protocols and counted using an EVE™ automatic cell counter (NanoEnTek). A cell solution of 100,000 cells/mL was prepared. For each 50 μL of transfection solution prepared, 400 μL of cell solution was prepared. The transfection solution was added to the cell solution and mixed. Note, the batch transfection can be scaled up to a final volume of 40 mL (transfection solution + cells). The cell + transfection solution was plated 30 μL per well using a small cassette for the Mutlidrop Combi Reagent Dispenser. To prevent striping, tubing was primed two times with the cell + transfection solution in the tubing for 2 min. Plates were centrifuged for 1 min at 1000 rpm before being incubated at 37 °C with 5% CO_2 for 16 h. At 16 h, plates were removed from the incubator and 10 μL of 40 μM compound (or 40% DMSO) in supplemented DMEM from an intermediate compound plate (see below) was added to each well with a Biomek liquid handler. Plates were returned to the incubator for an additional 8 hr. Plates were read at 24 h after initial plating by removing the media manually and smacking the plate on a wicking or absorbent pad. Using the

multidrop, 40 μL Opti-MEM™ followed by 10 μL NanoGlo Live Cell reagent (reconstituted according to manufacturer's recommendation) was added to each well. Plates were incubated at room temperature, covered for 30 min and the chemiluminescence was measured by an EnVision plate reader.

Preparation of Intermediate Compound Plate

An intermediate compound plate was prepared by adding 60 nL of 10 mM compound or DMSO to each well of a 384-well plate from a low dead volume plate using the ECHO liquid handler. Prior to addition to assay plates, 15 μL of warmed cell culture media (DMEM + 10% FBS, L-glutamine, hygromycin, and pen-strep) was added to each well. This protocol was repeated with different volumes of compound and DMSO for concentration response experiments to achieve the final concentrations listed in Table 5.1.

Table 5.1. Concentrations of compound tested in RiPCA CRC.

Concentration	[Final] (μM)
1	0.391
2	0.781
3	1.563
4	3.125
5	6.25
6	12.5
7	25
8	50

Assessment of Cell Viability:

Cell Titer Glo384-well Protocol

Flp-In-293 SmHT cells were plated using a small cassette with the multidrop after tubing was primed twice for 2 min each. Each well contained 3,000 cells in 30 μL. Plates were centrifuged at

1000 rpm for 1 min and incubated at 37ºC and 5% CO_2 for 16 h. At 16 h, 10 μL of 40 μM compound in media was transferred to each well using the Biomek liquid handler. Plates were returned to the incubator for an additional 8 hr. Plates were read at 24 h after initial plating by removing the media manually and smacking the plate on a wicking or absorbent pad. Using the multidrop, 25 μL Opti-MEM™ followed by 25 μL Cell Titer Glo reagent (reconstituted according to manufacturer's recommendation) was added to each well. Plates were mixed on a plate shaker for 2 min and then incubated at room temperature, covered for 10 min and the chemiluminescence was measured by a PHERAstar plate reader.

cat-ELCCA:

White high binding capacity streptavidin-coated 384-well plates were washed three times with 50 μL phosphate buffered saline (PBS; 100 mM phosphate, 150 mM NaCl, pH 7.0 diluted from 10X) using a Bio-Tek Elx405 plate washer. Lin28 was immobilized by adding 10 μL of 200 nM biotinylated Lin28 (see Lin28-HT-biotin protocol) in storage buffer (20 mM Tris pH 7.8, 100 mM. KCl, 0.2 mM EDTA, 10% glycerol (v/v), 0.005% Tween-20) to each well by multidrop. To prevent striping, tubing was primed two times with the Lin28 solution in the tubing for 2 min each. Plates were sealed with plate tape and centrifuged for 1 min at 1000 rpm before being stored at 4 °C overnight. Following the overnight incubation, plates were washed three times with 50 μL PBS and smacked against a spill pad to remove excess moisture. Compounds or DMSO were added to each well via ECHO liquid handler to achieve final concentrations listed in Table 5.2. Then 10 μL of 200 nM pre-let-7d-TCO in binding buffer (50 mM Tris pH 7.6, 150 mM NaCl, 5% glycerol, 0.05% Tween-20, freshly added 1 mM $ZnCl_2$ and 10 mM β-mercaptoethanol) was added to lanes 1-23 by multidrop. Lane 24 contained binding buffer only and was added manually with a

multichannel repeater pipette. Plates were centrifuged for 1 min at 1000 rpm and incubated for one hour at room temperature. After the incubation, wells were washed three times with 50 μL PBS and smacked against a spill pad to remove excess moisture. Next, 10 μL of 750 nM mTet-HRP in PBS was added to each well by multidrop and incubated for one hour at room temperature. Finally, wells were washed three times with 80 μL PBS-T (PBS with 0.05% Tween-20), followed by three times with 80 μL PBS. Chemiluminescence signal was measured after the addition of 25 μL of prepared SuperSignal West Pico reagent on a PHERAstar plate reader.

Table 5.2. Concentrations of compound tested in cat-ELCCA for CRC.

Concentration	[Final] (μM)
1	0.046
2	0.14
3	0.41
4	1.23
5	3.7
6	11.1
7	33.3
8	100

Dicer Digest Assay:

Pre-let-7d-4-Cl (500 nM in Arenz buffer; 20 mM Tris pH 7.6, 12 mM NaCl, 2.5 mM $MgCl_2$, freshly added 1 mM DTT) was pre-incubated with 50 μM compound (final concentration of DMSO of 5%) for 15 min at room temperature. Then 1 μL of 1 mg/mL Dicer was added, and reactions were incubated at 37 °C for 2 hr. Reactions were quenched with 2.5 μL of 80% glycerol and 6.25 μL was run on a 10% TBE-Urea gel at 200V for 30 min. Gels were stained with SYBR Gold for ~5 min and imaged with ChemiDoc™ Imaging System (Bio-Rad).

RT-qPCR:

T47d cells were plated in a 6-well plate at a density of 200,000 cells/mL and incubated for 24 h at 37 °C with 5% CO_2. Media was then removed and replaced with fresh media containing DMSO or 5 μM compound ([final] DMSO of 0.05%) and returned to the incubator for 48 hr. Small RNAs (<200 nt) were harvested with the *mir*Vana™ miRNA Isolation Kit (Thermo cat #AM1561) following the manufacturer's protocol for enrichment of small RNAs. cDNA was synthesized from 50 ng small RNA using the Taqman™ Advanced miRNA cDNA Synthesis Kit (Thermo cat #A28007) following the manufacturer's protocol. RT-qPCR was performed in 384-well plate with a 1:5 cDNA dilution following the Taqman™ Advanced miRNA Assay (Thermo cat #4444557) protocol using probes for miR-16-5p, miR-21-5p, let-7a-5p, let-7d-5p, and let-7g-5p (Thermo) on a QuantStudio™5 thermocycler (Thermo) using the fast Taqman™ $\Delta\Delta C_T$ protocol. Relative fold change was calculated using the comparative threshold cycle ($\Delta\Delta C_T$) method.

5.7 References

1. Roos, M.; Pradere, U.; Ngondo, R. P.; Behera, A.; Allegrini, S.; Civenni, G.; Zagalak, J. A.; Marchand, J. R.; Menzi, M.; Towbin, H.; Scheuermann, J.; Neri, D.; Caflisch, A.; Catapano, C. V.; Ciaudo, C.; Hall, J., A Small-Molecule Inhibitor of Lin28. *ACS Chem Biol* **2016,** *11* (10), 2773-2781.

2. Sin-Chan, P.; Mumal, I.; Suwal, T.; Ho, B.; Fan, X.; Singh, I.; Du, Y.; Lu, M.; Patel, N.; Torchia, J.; Popovski, D.; Fouladi, M.; Guilhamon, P.; Hansford, J. R.; Leary, S.; Hoffman, L. M.; Levy, J. M. M.; Lassaletta, A.; Solano-Paez, P.; Rivas, E.; Reddy, A.; Gillespie, G. Y.; Gupta, N.; Van Meter, T. E.; Nakamura, H.; Wong, T. T.; Ra, Y. S.; Kim, S. K.; Massimi, L.; Grundy, R. G.; Fangusaro, J.; Johnston, D.; Chan, J.; Lafay-Cousin, L.; Hwang, E. I.; Wang, Y.; Catchpoole, D.; Michaud, J.; Ellezam, B.; Ramanujachar, R.; Lindsay, H.; Taylor, M. D.; Hawkins, C. E.; Bouffet, E.; Jabado, N.; Singh, S. K.; Kleinman, C. L.; Barsyte-Lovejoy, D.; Li, X. N.; Dirks, P. B.; Lin,

C. Y.; Mack, S. C.; Rich, J. N.; Huang, A., A C19MC-LIN28A-MYCN Oncogenic Circuit Driven by Hijacked Super-enhancers Is a Distinct Therapeutic Vulnerability in ETMRs: A Lethal Brain Tumor. *Cancer Cell* **2019,** *36* (1), 51-+.

3. Lim, D.; Byun, W. G.; Koo, J. Y.; Park, H.; Park, S. B., Discovery of a Small-Molecule Inhibitor of Protein-MicroRNA Interaction Using Binding Assay with a Site-Specifically Labeled Lin28. *Journal of the American Chemical Society* **2016,** *138* (41), 13630-13638.

4. Lightfoot, H. L.; Miska, E. A.; Balasubramanian, S., Identification of small molecule inhibitors of the Lin28-mediated blockage of pre-let-7g processing. *Org Biomol Chem* **2016,** *14* (43), 10208-10216.

5. Wang, L.; Rowe, R. G.; Jaimes, A.; Yu, C.; Nam, Y.; Pearson, D. S.; Zhang, J.; Xie, X.; Marion, W.; Heffron, G. J.; Daley, G. Q.; Sliz, P., Small-Molecule Inhibitors Disrupt let-7 Oligouridylation and Release the Selective Blockade of let-7 Processing by LIN28. *Cell Rep* **2018,** *23* (10), 3091-3101.

6. Borgelt, L.; Li, F.; Hommen, P.; Lampe, P.; Hwang, J.; Goebel, G. L.; Sievers, S.; Wu, P., Trisubstituted Pyrrolinones as Small-Molecule Inhibitors Disrupting the Protein-RNA Interaction of LIN28 and Let-7. *Acs Med Chem Lett* **2021,** *12* (6), 893-898.

7. Lorenz, D. A.; Kaur, T.; Kerk, S. A.; Gallagher, E. E.; Sandoval, J.; Garner, A. L., Expansion of cat-ELCCA for the Discovery of Small Molecule Inhibitors of the Pre-let-7-Lin28 RNA-Protein Interaction. *Acs Med Chem Lett* **2018,** *9* (6), 517-521.

8. Bar, H.; Zweifach, A., Z' Does Not Need to Be > 0.5. *SLAS Discov* **2020,** *25* (9), 1000-1008.

9. Song, J. M.; Menon, A.; Mitchell, D. C.; Johnson, O. T.; Garner, A. L., High-Throughput Chemical Probing of Full-Length Protein-Protein Interactions. *ACS Comb Sci* **2017,** *19* (12), 763-769.

CHAPTER 6
Conclusions and Future Directions

The pivotal role that miRNAs play in practically all aspects of cellular biology and the contribution of dysregulated miRNAs to the development or progression of diseases has highlighted this class of RNA as a promising pathway towards novel therapeutics. Though large-scale studies have identified putative pre-miRNA binding proteins[1], much more work is required to elucidate the functional role of these interactions. Characterizing miRNA-protein interactions is crucial in understanding of mechanisms of miRNA regulation as well as identifying potential avenues for therapeutic intervention. The development of RPI detection assays has advanced our understanding of specific RPIs and enabled the discovery of inhibitors of such interactions, yet much of the available technology has not been or is unable to be applied to the study of RPIs involving highly processed RNAs. The present work details the development of a cell-based platform, RiPCA, for the characterization, validation, and manipulation of pre-miRNA-protein interactions. While there are some notable limitations, the RiPCA technology can be further developed and has the potential to be broadly applicable to the study of RPIs in live cells.

6.1 RiPCA: Limitations and Further Development

While RiPCA fills a key technological gap by providing a platform for the detection of RPIs in live cells, there are several elements that limit its functionality. First, RiPCA necessitates artificial expression of the RBP-of-interest, which could alter protein or RNA localization or

cellular processes unfavorably. Further, it is hypothesized that RiPCA signal with certain constructs, namely hnRNP A1, is limited by competition for binding to the RNA probe with endogenously expressed protein. Since excessive LgBiT expression can cause undesired background signal, there is a limit to the amount of RBP-LgBiT that can be expressed to compensate for competition.

One potential solution is the development of an alternative approach to labeling the RBP in the assay. In this version, cell lines in which the RBP-of-interest would be tagged at the endogenous locus with the LgBiT fragment would be generated with CRISPR.[2] These cell lines could then be transiently transfected with a SmBiT-HaloTag construct and RiPCA RNA probe. This system would more closely mirror RBP localization and expression levels, while maintaining the ability to tune S/B with modular SmHT expression. It is expected that this set-up would enable detection of RPIs involving highly expressed RBPs and, furthermore, could be performed in any cell line compatible with transfection, creating the potential to be used in a variety of disease contexts.

Additionally, RiPCA requires a chemically modification of the RNA probe. While there have not been significant challenges or perturbations of the systems tested, it is conceivable that other interactions might be disturbed by the presence of a PEG linker and HaloTag ligand in the RNA -of-interest. Additionally, as seen in Chapter 4, the position of the modification is important for the successful detection of signal. It might be required to screen several modified uridine positions in a given RNA probe in order to detect S/B in RiPCA. Notably, there are a few other modified amino acids that can be incorporated via chemical synthesis to label RNA probes on the 5’ or 3’ terminus of a sequence; however, there is no option for incorporating amino-modified adenosine, guanosine, or cytidine from commercial sources.[3] Other options for nucleic acid

conjugation to the HaloTag ligand should be explored to expand the sites on an RNA-of-interest that can be labeled for RiPCA.

Finally, given that RiPCA is a cell-based assay in which inhibition is indicated by loss of signal, the false hit rate could be quite high depending on the average cytotoxicity of compounds in the screened library. To mitigate the selection of false positive hits, it is advisable to factor in cell viability data when selecting hits in RiPCA. This could be done by pre-screening for cytotoxicity and either eliminating exceedingly toxic compounds or factoring in the percentage of cell death occurring when calculating the percent inhibition in RiPCA.

6.2 Broad Application of RiPCA for the Detection of RPIs

While this dissertation has focused on the detection and manipulation of pre-miRNA-protein interactions, understanding, characterizing, and manipulating RPIs involving other classes of RNA is of great interest to the scientific community. The general applicability of RiPCA in detecting RPIs involving pre-miRNA hairpins and several RBPs was demonstrated in Chapter 4. Coupled with the successful adaptation of RiPCA to HTS as reported in Chapter 5, this novel platform promises to be a powerful tool for broader RPI detection and the advancement of efforts in RNA-targeted drug discovery. Efforts are under way in the Garner laboratory to adapt RiPCA for the detection of RPIs involving mRNAs and lncRNAs as well as RBPs that function as RNA modification readers.

6.3 The Future of RNA-targeted Small Molecule Drug Discovery

The FDA approval of risdiplam, a small molecule that directly binds a stem-loop structure in the *SMN2* pre-mRNA, has encouraged the field to expand efforts to discover RNA-binding therapeutics.[4] Chapter 5 presented the discovery of small molecules capable of inhibiting the pre-

let-7d/Lin28 interactions in cells (RiPCA) and *in vitro* (cat-ELCCA). Unfortunately, these compounds demonstrated limited activity in cells. While the goal of limiting the library of compounds screened to reported or predicted RNA binding molecules was to increase chances of successfully identifying an RNA-targeting inhibitor, relatively few chemical moieties that demonstrate RNA-specific binding are known. Perhaps more success would be seen if the amount of chemical space represented in the molecules screened were to be expanded. Further miniaturization of RiPCA to 1536-well format would similarly enhance small molecule discovery effort and given the number of cells utilized in other cell-based assays in 1536-well format[5], it is likely that RiPCA would be amenable to this format.

6.4 Alternative Therapeutic Strategies for Targeting RNAs

Several recent scientific advances have opened avenues for the targeting of RNA biology through unconventional mechanisms. First, the success of mRNA vaccines in combating the pandemic caused by the SARS-CoV-2 virus demonstrates the utility of protein-based therapeutics.[6] Second, the success of technologies designed for the directed evolution of proteins has allowed researchers to achieve highly specific, unnatural activities.[7] One could envision that in combination, these strategies could eventually lead to the development of protein-based medicines to target RNAs, whether through their processing, abundance, or activity.

6.5 Concluding Remarks

Mapping and characterizing the functional consequences of RPIs will not only complete our understanding of biological processes and regulation but may also present a vast pool of targets for novel approaches to treating human diseases. This has necessitated the development of tools to enable the study and modulation of these interactions. Limitations notwithstanding, RiPCA

promises to serve as a useful live-cell platform for the detection, validation, and manipulation of RPIs as well as a solid foundation for further optimization and assay development.

6.6 References

1. Treiber, T.; Treiber, N.; Plessmann, U.; Harlander, S.; Daiss, J. L.; Eichner, N.; Lehmann, G.; Schall, K.; Urlaub, H.; Meister, G., A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Molecular Cell* **2017,** *66* (2), 270-284.

2. Xiang, X.; Li, C.; Chen, X.; Dou, H.; Li, Y.; Zhang, X.; Luo, Y., CRISPR/Cas9-Mediated Gene Tagging: A Step-by-Step Protocol. *Methods Mol Biol* **2019,** *1961*, 255-269.

3. Rosenblum, S. L.; Garner, A. L., RiPCA: An Assay for the Detection of RNA-Protein Interactions in Live Cells. *Curr Protoc* **2022,** *2* (2), e358.

4. Sheridan, C., First small-molecule drug targeting RNA gains momentum. *Nat Biotechnol* **2021,** *39* (1), 6-8.

5. Knight, S.; Plant, H.; McWilliams, L.; Murray, D.; Dixon-Steele, R.; Varghese, A.; Harper, P.; Ramne, A.; McArdle, P.; Engberg, S.; Bennett, N.; Blackett, C.; Wigglesworth, M., Enabling 1536-Well High-Throughput Cell-Based Screening through the Application of Novel Centrifugal Plate Washing. *Slas Discovery* **2017,** *22* (6), 732-742.

6. Hogan, M. J.; Pardi, N., mRNA Vaccines in the COVID-19 Pandemic and Beyond. *Annu Rev Med* **2022,** *73*, 17-39.

7. McLure, R. J.; Radford, S. E.; Brockwell, D. J., High-throughput directed evolution: a golden era for protein science. *Trends in Chemistry* **2022,** *4* (5), 378-391.