Facilitating the Discovery of Chemical Tools for Manipulating pre-miRNA Structure-Function

By

Daniel A. Lorenz

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2018

Doctoral Committee:

Assistant Professor Amanda L. Garner, Chair
Associate Professor Tomasz Cierpicki
Assistant Professor Jayakrishnan Nandakumar
Associate Professor Patrick J. O'Brien

Daniel A. Lorenz

Lorenzda@umich.edu

ORCID iD: 0000-0002-9249-7338

Acknowledgments

I would first like to thank Professor Amanda Garner for taking me under her wing and fostering my growth as a scientist. Her dedication and passion to both her science and students has been inspirational. I would also like to thank my committee, the Program in Chemical Biology, and the University of Michigan for providing me an opportunity and creating a strong foundation for my future. A special thank you to my undergraduate mentors Professors Jon Fukuto and Joseph Lin for sparking my interest in chemical biology.

I am forever grateful to the many family and friends that never gave up on me and continue to support me. In particular, to my mother, you have been my role model. You have tackled life challenges head on and never let anything stop me from achieving dreams. To my rocks at home, Sabrina and Maddie, thank you for never failing to put a smile on my face. Finally, to my fathers in the sky, thank you for walking with me every step of the way.

Table of Contents

List of Figures

List of Tables

Abstract

The biological implications of RNA continue to expand beyond the role as the simple messenger between DNA and protein translation. This expansion has been catalyzed by the development of new technologies that facilitate both the discovery of new RNAs and their biological impact. One class of RNA that has been particularly interesting is the family of small non-coding RNAs termed microRNAs (miRNAs). While miRNAs do not code for proteins, their biology is intimately intertwined, as miRNAs are estimated to regulate the majority of protein translation. This translational regulation stems from a miRNA's ability to use sequence complementarity to identify a target messenger RNA and trigger translational suppression.

Since miRNAs play crucial roles in biology, their dysregulation has been implicated in nearly every disease. As such, miRNAs are being evaluated as targets for therapeutic intervention. There are many different strategies being explored to manipulate these small RNAs. The furthest progressed strategy for miRNA manipulation is based on modified oligonucleotides and is being evaluated in the clinic; however, drugs developed using this approach have yet to receive FDA approval. While oligonucleotides are easy to design, obtaining optimal drug like properties has remained a challenge. Alternative strategies for miRNA manipulation, such as small molecules, have remained at the proof-of-concept stage largely due to the challenges associated with their discovery and design.

To facilitate the discovery of new small molecules and strategies to manipulate miRNA biology, new technologies were developed. These reported techniques focused on the generation of RNA-protein conjugates that were crucial to the development of both high-throughput screening and biological assays. The first of these assays is focused on inhibiting the biogenesis of miRNAs and was able to identify natural product extracts as a source of small molecules capable of binding and disrupting miRNA precursors. The second assay platform discovered a small molecule capable of disrupting a known miRNA precursor-RNA binding protein interaction, pre-let-7d-Lin28. Having successfully disrupted a clinically relevant RNA-protein interaction, a final assay was created to aid in the discovery of new RNA-protein interactions by reporting direct interactions in live cells. Altogether, the technologies reported serve as a launching platform for the discovery of molecules with the ability to manipulate clinically significant miRNAs.

Chapter 1

microRNA Biological Significance and Biogenesis: An Overview

microRNAs (miRNAs) were first discovered in 1993 as small non-coding RNAs that use sequence complementarity to regulate gene expression post transcriptionally.[1,2] Yet, the impact of these small RNAs in biology remained underestimated until nearly 10 years later when miRNAs were found to be conserved across numerous classes in the animal kingdom.[3] In the two decades to follow, the role of miRNAs in biology exploded, quickly revealing their role in gene regulation. With the completion of the human genome project, development of high-throughput transcriptome sequencing, and expanded bioinformatic tools, over 2,500 miRNAs have been discovered in *Homo sapiens*.[4-8] This ever-expanding class of RNA is predicted to regulate more than 60% of protein coding transcripts, establishing miRNAs as indispensable for proper cellular regulation.[9]

1.1 Biogenesis

With few exceptions, miRNAs follow a canonical biogenesis pathway that is similarly conserved across species (Figure 1.1).[10] miRNAs are encoded at locations throughout the genome and can be localized in clusters, inside other genes, or as unique transcriptional sites that are transcribed by RNA Polymerase II and termed pri-miRNAs.[11] These pri-miRNAs can vary in length, but contain a hairpin loop that is recognized by the

microprocessor complex which contains the core proteins DiGeorge Syndrome Critical Region 8 (DGCR8) and the RNase III enzyme Drosha.[12] Drosha is responsible for cutting at the base of the hairpin loop to generate ~60-80 nucleotide pre-miRNAs with a characteristic 2 nucleotide 3' overhang. pre-miRNAs are then exported to the cytoplasm through a Ran GTPase, Exportin 5, where they are recognized by another RNase III enzyme, Dicer.[13] Generally thought to be accompanied by accessory proteins, Dicer is responsible for the cleavage of the loop generating a ~22 nucleotide RNA duplex with overhangs on both 3' ends.[14] The strand with less thermodynamic stability at the 5' end is then transferred to an Argonaute (Ago) protein which forms the RNA Induced Silencing Complex (RISC).[15] Utilizing nucleotide positions 2-8 from the 5' end, termed the seed region, RISC finds complementary transcripts and causes translational suppression by sequestering the transcript from the translational machinery and recruiting deadenylation and decapping factors to reduce transcript stability.[16]
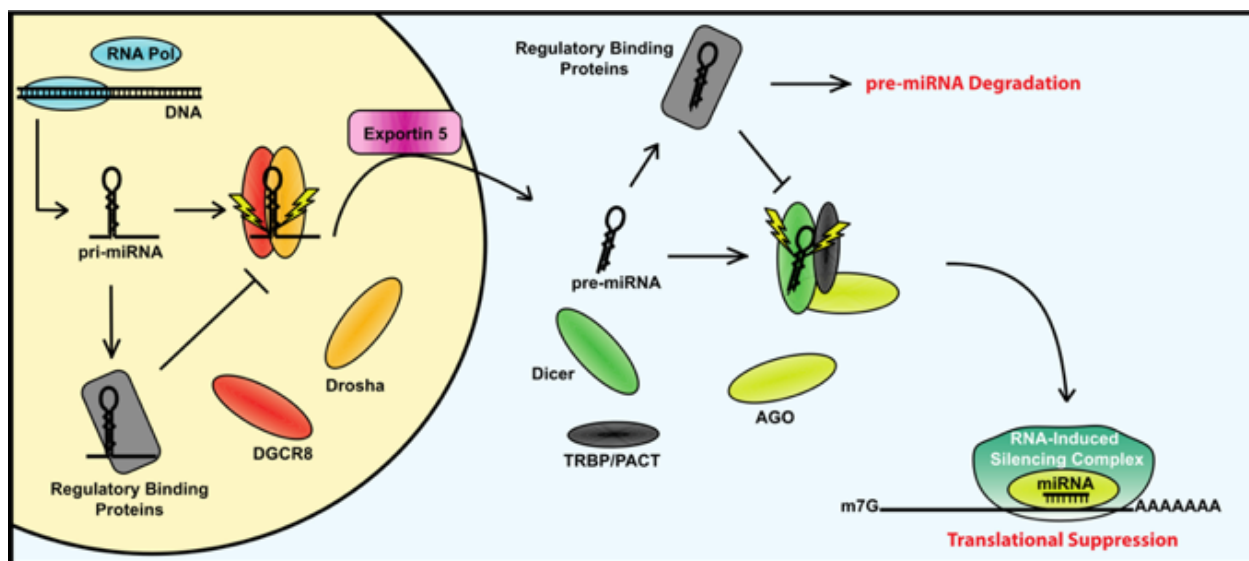
Figure 1.1. Canonical miRNA Biogenesis Pathway.

1.2 Dicer

Dicer is essential for the final step in miRNA biogenesis and continues to be the focus of numerous research efforts. This large ~220 kDa protein is composed of several domains including a helicase domain, a PIWI, AGO, and Zwille (PAZ) domain, two RNase III domains, a double-stranded RNA-binding domain, and another RNA-binding domain of unknown function (DUF).[17] To date, there is no atomic resolution structure of full-length human Dicer. Therefore, the current model for pre-miRNA recognition and processing by Dicer is based on the modeling of crystallized domains to electron microscopy results and the crystal structure of Giardia Dicer (Figure 1.2).[17,18] Based on these structural insights, Dicer is thought to act as a molecular ruler, binding the two nucleotide 3' overhang of a pre-miRNA with its PAZ domain and loop region with the helicase domain. This binding orients the two magnesium-dependent RNase III domains approximately 22 nucleotides from the PAZ domain to generate a second two nucleotide 3' overhang. While canonical miRNAs are thought to be processed by Dicer in this mechanism, every pre-miRNA has a unique sequence, length, and loop size, all of which have been shown to affect Dicer kinetics. Studies evaluating these differences have revealed that larger non-base-paired loop regions are processed faster than smaller loops, and both are processed faster than perfectly base-paired strands.[19,20] Additionally, the presence of Dicer-interacting proteins, such as TRBP and PACT, have been shown to effect Dicer cleavage rates and mature miRNA levels.[21,22]

Figure 1.2. Giardia Dicer Crystal Structure. (PDB:2FFL) overlaid with a modeled pre-let-7d substrate (MC-Fold). Dicer domains are colored cyan (PAZ), magenta and yellow (RNase III domains) with magnesium ions colored red.

1.3 miRNAs Role in Disease

Altered miRNA expression patterns have been observed between normal and pathological tissues.[23] Aided by advancements in microarray technology, miRNAs profiling has become a standard practice with the capacity to analyze hundreds of miRNAs at a time.[24] However, these arrays simply identify miRNAs with altered expression levels and require subsequent follow-up to determine if the miRNAs are sufficient to drive the disease. This validation is typically done through a combination of

knockout and overexpression experiments. Although, genetic manipulation of miRNAs does require careful analysis because many miRNAs display a high degree of tissue selectivity.[25] Nonetheless, the importance of miRNAs is highlighted by the embryonic lethality associated with the global inhibition of miRNA biogenesis resulting from deleting either Dicer or Drosha.[26] While some mutations of these proteins that cause heterozygous expression or inactivation are viable, these alterations are associated with pathological tissues.[27] With such far reaching regulatory abilities, many miRNAs have shown causative relationships with diseases such as cancer, viral infection, and cardiovascular disease, and also exhibit lethality when knocked out in specific tissues. Table 1 provides some of examples the more well characterized miRNAs and their associated diseases.[28-47]

Table 1.1 Examples of disease relevant miRNAs

miRNA	Target Gene	Expression Level	Disease	Reference
let-7	RAS	Down	Cancer	28
1	MET	Down	Cancer	29
10b	HOXD10	Up	Cancer	30
21	PTEN	Up	Cancer	31
27a	ZBTB10	Up	Cancer	32
29a	LOXL2	Down	Cancer	33
34a	BCL2	Down	Cancer	34
96	FOXO1	Up	Cancer	35
122	HCV, CCNG1	Up, Down	HCV, Cancer	36,37
125	EIF4EBP1	Down	Cancer	38
133a	CDC42	Down	Heart Disease	39
142	APC	Up	Cancer	40
155	TP53INP1	Up	Cancer	41
206	HDAC4	Up	ALS	42
335	RB1	Up	Cancer	43
372	LATS2	Up	Cancer	44
373	LATS2	Up	Cancer	44
504	P53	Up	Cancer	45
525	ZNF395	Up	Cancer	46
544	MTOR	Up	Cancer	47

1.4 Lin28-Let7

The miRNA let-7 differs from the traditional miRNA nomenclature as it was initially characterized as lethal gene 7 in *C. elegans*.[48] One of the first miRNAs discovered, let-7 was found to be indispensable for larval development. Let-7 was later identified as the first miRNA conserved across species, including *Homo sapiens*, and sparked the discovery of other human miRNAs.[3] Further genomic analysis revealed that there are 10 different mature let-7 family member sequences in humans.[28] These family members are all associated with cell differentiation and target key oncogenic proteins like Ras, HMGA2, and c-Myc.[49-51] Since these oncogenes are common drivers of tumor growth and progression, let-7 expression levels have been similarly linked to cancer and have been used as biomarkers for patient prognosis with low levels associated with poor patient prognosis.[52]

Cancer cells have developed multiple mechanisms to suppress let-7 expression and activity to encourage an undifferentiated cell state. The first is transcriptional, similar to other genes, let-7 transcription can be described by the methylation status of DNA with hypermethylated let-7 transcriptional sites associated with reduced transcription.[53] A second mechanism to decrease mature let-7 levels is to interfere post-transcriptionally with the miRNA biogenesis pathway, preventing mature let-7 from being produced.[54] Finally, the tumor suppressive properties of miRNAs can be negated by single nucleotide mutations (SNM).[55] These SNM are located in the miRNA complementary site on the target mRNA preventing RISC identification and subsequent translational suppression.

One of the most studied RNA binding proteins (RBP) that can block miRNA maturation post-transcriptionally is Lin28.[54] Lin28 serves as a master regulator of cell differentiation, and in addition to the ability to regulate many other transcripts, is capable

of suppressing mature let-7 levels.[56] The importance of Lin28 was highlighted when exogenous expression of Lin28 along with Oct4, Sox2, and Nanog was used to generate induced pluripotent stem cells from somatic cells.[57] Since Lin28 promotes an undifferentiated cell state and downregulates the tumor suppressive miRNA let-7 family, it is often overexpressed in cancer.[52] Overexpression of either of the two human homologues, Lin28A and B, can trigger reduced let-7 levels.[58] The main difference between Lin28A and B is their cellular localization, as they primarily localize to the cytoplasm and nucleus, respectively.[59] Both Lin28 proteins contain two RNA-binding sites, a cold shock domain (CSD) and two zinc knuckle domains (ZKD) that recognize a GNGAY and GGAG motif, respectively.[60] Crosslinking and immunoprecipitation coupled with high-throughput sequencing (CLIP-Seq) revealed that greater than 6,000 RNA transcripts contain a GGAG motif that interacts directly with Lin28.[56] This GGAG motif is present in the loop of pri and pre-let-7 family members. In addition to Lin28 regulating let-7, let-7 has been shown to directly regulate Lin28 levels through canonical miRNA mediated gene silencing creating an important negative feedback loop between Lin28 and let-7.[61] This double-negative feedback loop is believed to be an essential developmental switch in many organisms.[62]

Many studies have evaluated how Lin28 interacts with and regulates let-7 biogenesis. Nam et al. solved the first crystal structure of Lin28A in complex with let-7 substrates.[63] These structures, highlighted in Figure 1.3, are domain swapped dimers with one let-7 hidden for simplicity. Upon examining the structure, it was revealed that the CSD and ZKD bind the hairpin loop independently and are spaced apart by a flexible linker. However, many studies report the need for the CSD to bind first to remodel the

loop and allow for the ZKD to bind.[64,65] This bidentate binding mechanism is thought to contribute to specificity and increase the affinity ranging from high nanomolar to low micromolar depending on assay conditions. This high affinity complex is believed to bind both the pri and pre-let-7 miRNAs and block Drosha and Dicer maturation by sequestration.[28] In addition to the RNA-binding function of the ZKDs, they are also involved in the recruitment of Tutases that polyuridylate the 3' end of pre-let7 miRNAs, which signals the RNA for degradation by Dis3L2.[66]
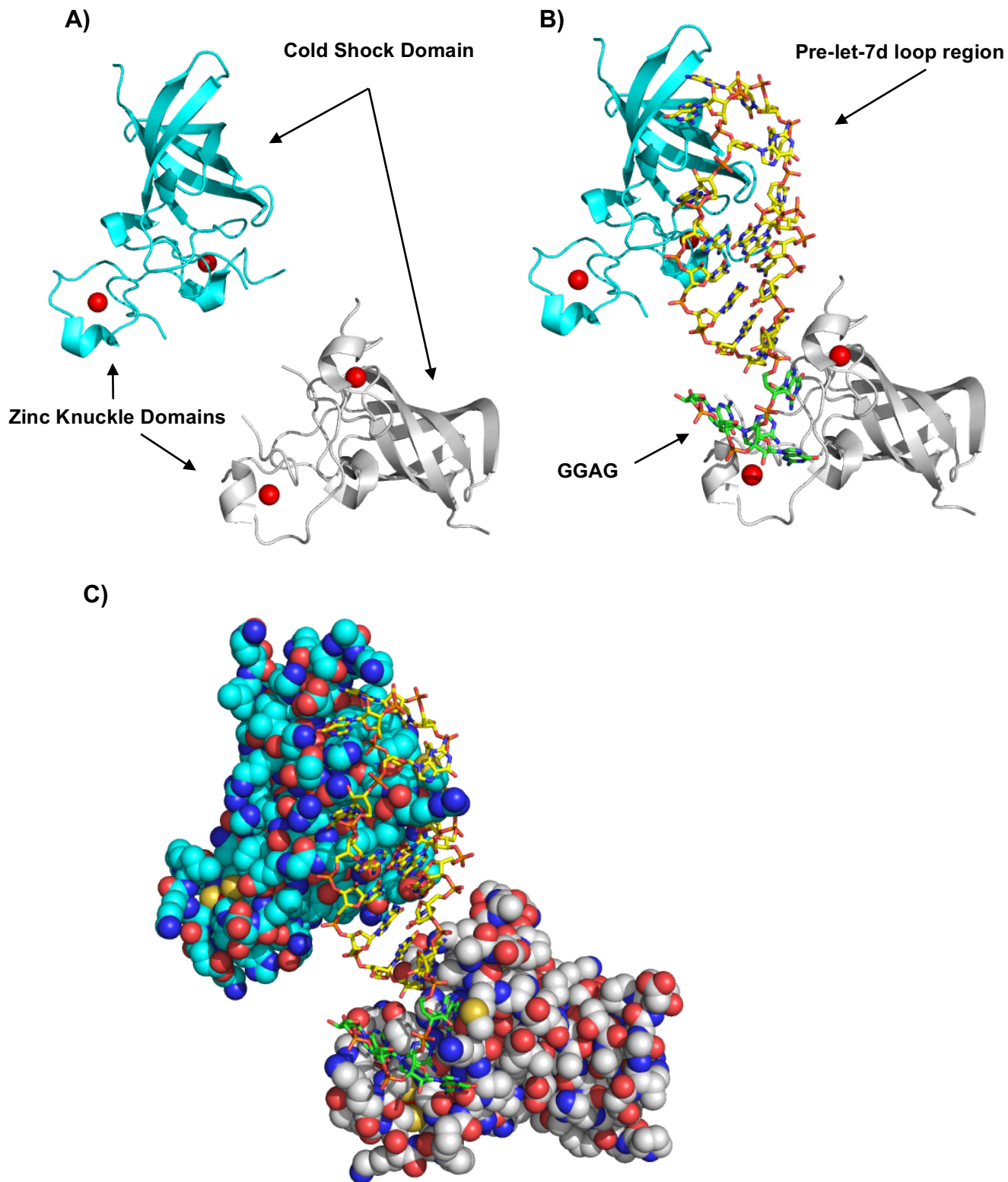
Figure 1.3. Lin28 Crystal Structure. (PDB: 3TRZ) A) with the zinc knuckle and cold shock domains labeled and RNA hidden B) one pre-let-7d shown with the GGAG motif highlighted in green C) Same depiction as B but atoms shown as spheres to show side chain arrangement

1.5 miR-21

There are also many examples of miRNAs that are overexpressed in disease; however, miR-21 exemplifies how influential a single miRNA can be. miR-21 was one of the first oncogenic miRNAs, or oncomiRs, described due to its high expression in lymphomas.[67] This overexpression was later characterized in breast, colon, lung, pancreatic, prostate, and stomach cancers, making it of the most intriguing and well-studied oncomiRs to date.[68] Targets of miR-21 include the tumor suppressor proteins PTEN and PDCD4, further suggesting that miR-21 overexpression is key for cancer progression.[69,70] One important question that remained was if miR-21 was causative or correlative. Slack and colleagues addressed this question by demonstrating a causative relationship between the overexpression of miR-21 in mice and the generation of pre-B-cell lymphomas.[71] These lymphomas were addicted to miR-21 expression and would return to normal upon the removal of miR-21 overexpression. This landmark study showcased that a single miRNA can drive cancer progression.

1.6 Concluding Remarks

Due to the high conservation of miRNAs across species, it comes as no surprise that miRNAs have been implicated in nearly every aspect of biology. As such, disruption of miRNA biogenesis, whether global or on an individual basis, has been equally involved in disease progression. Two miRNAs of particular interest, miR-21 and let-7, play crucial roles in biology and are currently being evaluated as therapeutic targets. Additionally, Dicer and Lin28 are also prime candidates for therapeutic interventions for their crucial role in several pathologies.

1.7 Copyright

The work in this chapter was reproduced in part from Lorenz, D. A. & Garner, A. L. in *RNA Therapeutics* (ed. Garner, A. L.) 79–110 (Springer International Publishing, 2018).

1.8 References

1. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75,** 843–854 (1993).

2. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75,** 855–862 (1993).

3. Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E. & Ruvkun, G. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408,** 86–89 (2000).

4. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Research* **32,** D109–11 (2004).

5. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A. & Enright, A. J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* **34,** D140–4 (2006).

6. Griffiths-Jones, S., Saini, H. K., van Dongen, S. & Enright, A. J. miRBase: tools for microRNA genomics. *Nucleic Acids Research* **36,** D154–8 (2008).

7. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research* **39,** D152–D157 (2010).

8. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42,** D68–73 (2014).

9. Friedman, R. C., Farh, K. K. H., Burge, C. B. & Bartel, D. P. Most mammalian

mRNAs are conserved targets of microRNAs. *Genome Research* **19,** 92–105 (2009).

10. Lin, S. & Gregory, R. I. MicroRNA biogenesis pathways in cancer. *Nature Publishing Group* **15,** 321–333 (2015).

11. 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.* **23,** 4051–4060 (2004).

12. Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S. Y., Cho, Y., Zhang, B.-T. & Kim, V. N. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125,** 887–901 (2006).

13. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363–366 (2001).

14. Tian, Y., Simanshu, D. K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V. N. & Patel, D. J. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular Cell* **53,** 606–616 (2014).

15. Hibio, N., Hino, K., Shimizu, E., Nagata, Y. & Ui-Tei, K. Stability of miRNA 5'terminal and seed regions is correlated with experimentally observed miRNA-mediated silencing efficacy. *Sci. Rep.* **2,** 996 (2012).

16. Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. & Izaurralde, E. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development* **20,** 1885–1898 (2006).

17. Lau, P.-W., Guiley, K. Z., De, N., Potter, C. S., Carragher, B. & MacRae, I. J. The molecular architecture of human Dicer. *Nature Structural & Molecular Biology* **19,** 436–440 (2012).

18. MacRae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P. D. & Doudna, J. A. Structural basis for double-stranded RNA processing by Dicer. *Science* **311,** 195–198 (2006).

19. Tsutsumi, A., Kawamata, T., Izumi, N., Tomari, Y. & Seitz, H. Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nature Structural & Molecular Biology* **18,** 1153–1158 (2011).

20. Koh, H. R., Ghanbariniaki, A. & Myong, S. RNA stem structure governs coupling of dicing and gene silencing in RNA interference. *Proceedings of the National Academy of Sciences* **114,** E10349–E10358 (2017).

21. Chakravarthy, S., Sternberg, S. H., Kellenberger, C. A. & Doudna, J. A. Substrate-

specific kinetics of Dicer-catalyzed RNA processing. *Journal of molecular biology* **404,** 392–402 (2010).

22. Noland, C. L., Ma, E. & Doudna, J. A. siRNA Repositioning for Guide Strand Selection by Human Dicer Complexes. *Molecular Cell* **43,** 110–121 (2011).

23. Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R. & Golub, T. R. MicroRNA expression profiles classify human cancers. *Nature Publishing Group* **435,** 834–838 (2005).

24. Liu, C.-G., Calin, G. A., Volinia, S. & Croce, C. M. MicroRNA expression profiling using microarrays. *Nature protocols* **3,** 563–578 (2008).

25. Ludwig, N., Leidinger, P., Becker, K., Backes, C., Fehlmann, T., Pallasch, C., Rheinheimer, S., Meder, B., Stähler, C., Meese, E. & Keller, A. Distribution of miRNA expression across human tissues. *Nucleic Acids Research* **44,** 3865–3877 (2016).

26. Park, C. Y., Choi, Y. S. & McManus, M. T. Analysis of microRNA knockouts in mice. *Hum. Mol. Genet.* **19,** R169–75 (2010).

27. Foulkes, W. D., Priest, J. R. & Duchaine, T. F. DICER1: mutations, microRNAs and mechanisms. *Nature reviews. Cancer* **14,** 662–672 (2014).

28. Roush, S. & Slack, F. J. The let-7 family of microRNAs. *Trends in Cell Biology* **18,** 505–516 (2008).

29. Han, C., Yu, Z., Duan, Z. & Kan, Q. Role of MicroRNA-1 in Human Cancer and Its Therapeutic Potentials. *BioMed Research International* **2014,** 1–11 (2014).

30. Ma, L. Role of miR-10b in breast cancer metastasis. *Breast Cancer Res.* **12,** 210 (2010).

31. Pfeffer, S. R., Yang, C. H. & Pfeffer, L. M. The Role of miR-21 in Cancer. *Drug Dev. Res.* **76,** 270–277 (2015).

32. Chhabra, R., Dubey, R. & Saini, N. Cooperative and individualistic functions of the microRNAs in the miR-23a~27a~24-2 cluster and its implication in human diseases. *Molecular cancer* **9,** 232 (2010).

33. Wang, Y., Zhang, X., Li, H., Yu, J. & Ren, X. The role of miRNA-29 family in cancer. *European Journal of Cell Biology* **92,** 123–128 (2013).

34. Hermeking, H. The miR-34 family in cancer and apoptosis. *Cell Death Differ.* **17,** 193–199 (2010).

35. Guttilla, I. K. & White, B. A. Coordinate regulation of FOXO1 by miR-27a, miR-96, and miR-182 in breast cancer cells. *Journal of Biological Chemistry* **284,** 23204–23216 (2009).

36. Henke, J. I., Goergen, D., Zheng, J., Song, Y., Schüttler, C. G., Fehr, C., Jünemann, C. & Niepmann, M. microRNA-122 stimulates translation of hepatitis C virus RNA. *The EMBO journal* **27,** 3300–3310 (2008).

37. Gramantieri, L., Ferracin, M., Fornari, F., Veronese, A., Sabbioni, S., Liu, C.-G., Calin, G. A., Giovannini, C., Ferrazzi, E., Grazi, G. L., Croce, C. M., Bolondi, L. & Negrini, M. Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma. *Cancer Research* **67,** 6092–6099 (2007).

38. Ozen, M., Creighton, C. J., Ozdemir, M. & Ittmann, M. Widespread deregulation of microRNA expression in human prostate cancer. *Oncogene* **27,** 1788–1793 (2008).

39. Thum, T., Catalucci, D. & Bauersachs, J. MicroRNAs: novel regulators in cardiac development and disease. *Cardiovasc. Res.* **79,** 562–570 (2008).

40. Isobe, T., Hisamori, S., Hogan, D. J., Zabala, M., Hendrickson, D. G., Dalerba, P., Cai, S., Scheeren, F., Kuo, A. H., Sikandar, S. S., Lam, J. S., Qian, D., Dirbas, F. M., Somlo, G., Lao, K., Brown, P. O., Clarke, M. F. & Shimono, Y. miR-142 regulates the tumorigenicity of human breast cancer stem cells through the canonical WNT signaling pathway. *eLife* **3,** 3983–23 (2014).

41. Faraoni, I., Antonetti, F. R., Cardone, J. & Bonmassar, E. miR-155 gene: a typical multifunctional microRNA. *Biochim. Biophys. Acta* **1792,** 497–505 (2009).

42. Williams, A. H., Valdez, G., Moresi, V., Qi, X., McAnally, J., Elliott, J. L., Bassel-Duby, R., Sanes, J. R. & Olson, E. N. MicroRNA-206 delays ALS progression and promotes regeneration of neuromuscular synapses in mice. *Science* **326,** 1549–1554 (2009).

43. Scarola, M., Schoeftner, S., Schneider, C. & Benetti, R. miR-335 Directly Targets Rb1 (pRb/p105) in a Proximal Connection to p53-Dependent Stress Response. *Cancer Research* **70,** 6925–6933 (2010).

44. Staedel, C., Varon, C., Nguyen, P. H., Vialet, B., Chambonnier, L., Rousseau, B., Soubeyran, I., Evrard, S., Couillaud, F. & Darfeuille, F. Inhibition of Gastric Tumor Cell Growth Using Seed-targeting LNA as Specific, Long-lasting MicroRNA Inhibitors. *Mol Ther Nucleic Acids* **4,** e246 (2015).

45. Hu, W., Chan, C. S., Wu, R., Zhang, C., Sun, Y., Song, J. S., Tang, L. H., Levine, A. J. & Feng, Z. Negative Regulation of Tumor Suppressor p53 by MicroRNA miR-504. *Molecular Cell* **38,** 689–699 (2010).

46. Pang, F., Zha, R., Zhao, Y., Wang, Q., Chen, D., Zhang, Z., Chen, T., Yao, M., Gu, J. & He, X. MiR-525-3p Enhances the Migration and Invasion of Liver Cancer Cells by Downregulating ZNF395. *PLoS ONE* **9,** e90867–8 (2014).

47. Haga, C. L., Velagapudi, S. P., Strivelli, J. R., Yang, W.-Y., Disney, M. D. & Phinney, D. G. Small Molecule Inhibition of miR-544 Biogenesis Disrupts Adaptive Responses to Hypoxia by Modulating ATM-mTOR Signaling. *ACS chemical biology* **10,** 2267–2276 (2015).

48. 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* **403,** 901–906 (2000).

49. 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* **120,** 635–647 (2005).

50. Mayr, C., Hemann, M. T. & Bartel, D. P. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315,** 1576–1579 (2007).

51. 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 Research* **67,** 9762–9770 (2007).

52. Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T. & Takahashi, T. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Research* **64,** 3753–3756 (2004).

53. Lu, L., Katsaros, D., la Longrais, de, I. A. R., Sochirca, O. & Yu, H. Hypermethylation of let-7a-3 in epithelial ovarian cancer is associated with low insulin-like growth factor-II expression and favorable prognosis. *Cancer Research* **67,** 10117–10122 (2007).

54. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320,** 97–100 (2008).

55. Chin, L. J., Ratner, E., Leng, S., Zhai, R., Nallur, S., Babar, I., Müller, R.-U., Straka, E., Su, L., Burki, E. A., Crowell, R. E., Patel, R., Kulkarni, T., Homer, R., Zelterman, D., Kidd, K. K., Zhu, Y., Christiani, D. C., Belinsky, S. A., Slack, F. J. & Weidhaas, J. B. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Research* **68,** 8535–8540 (2008).

56. Wilbert, M. L., Huelga, S. C., Kapeli, K., Stark, T. J., Liang, T. Y., Chen, S. X., Yan, B. Y., Nathanson, J. L., Hutt, K. R., Lovci, M. T., Kazan, H., Vu, A. Q., Massirer, K. B., Morris, Q., Hoon, S. & Yeo, G. W. LIN28 Binds Messenger RNAs at GGAGA Motifs and Regulates Splicing Factor Abundance. *Molecular Cell* **48,** 195–206 (2012).

57. Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I. & Thomson, J. A. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318,** 1917–1920 (2007).

58. 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* **147,** 1066–1079 (2011).

59. Zhou, J., Ng, S.-B. & Chng, W.-J. LIN28/LIN28B: an emerging oncogenic driver in cancer stem cells. *Int. J. Biochem. Cell Biol.* **45,** 973–978 (2013).

60. Balzeau, J., Balzeau, J., Menezes, M. R., Menezes, M. R., Cao, S., Cao, S. & Hagan, J. P. The LIN28/let-7 Pathway in Cancer. *Front Genet* **8,** 31 (2017).

61. 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. *Nature cell biology* **10,** 987–993 (2008).

62. Rehfeld, F., Rohde, A. M., Nguyen, D. T. T. & Wulczyn, F. G. Lin28 and let-7: ancient milestones on the road from pluripotency to neurogenesis. *Cell Tissue Res* **359,** 145–160 (2015).

63. Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. *Cell* **147,** 1080–1091 (2011).

64. Mayr, F., Schütz, A., Döge, N. & Heinemann, U. The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Research* **40,** 7492–7506 (2012).

65. Desjardins, A., Bouvette, J. & Legault, P. Stepwise assembly of multiple Lin28 proteins on the terminal loop of let-7 miRNA precursors. *Nucleic Acids Research* **42,** 4615–4628 (2014).

66. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nature Structural & Molecular Biology* **16,** 1021–1025 (2009).

67. Iorio, M. V., Ferracin, M., Liu, C.-G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Ménard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M. & Croce,

C. M. MicroRNA gene expression deregulation in human breast cancer. *Cancer Research* **65,** 7065–7070 (2005).

68.   Volinia, S., Calin, G. A., Liu, C.-G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C. & Croce, C. M. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences* **103,** 2257–2261 (2006).

69.   Meng, F., Henson, R., Wehbe Janek, H., Ghoshal, K., Jacob, S. T. & Patel, T. MicroRNA-21 Regulates Expression of the PTEN Tumor Suppressor Gene in Human Hepatocellular Cancer. *Gastroenterology* **133,** 647–658 (2007).

70.   Frankel, L. B., Christoffersen, N. R., Jacobsen, A., Lindow, M., Krogh, A. & Lund, A. H. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *Journal of Biological Chemistry* **283,** 1026–1033 (2008).

71.   Medina, P. P., Nolde, M. & Slack, F. J. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature* **467,** 86–90 (2010).

Chapter 2

Current RNA Therapeutic Strategies: An Overview

The strategy of therapeutically targeting RNA has come to the forefront of biomedical research due to the emerging roles that RNAs play in nearly every disease. As highlighted in Chapter 1, miRNAs have an essential role in regulating biology and can drive a myriad of disease states. However, no miRNA therapies have received FDA approval. In fact, this trend extends to all RNA, as roughly 85% of FDA approved therapies target proteins, despite only representing 2% of the genome.[1] To target the other 90% of our genome that is thought to be transcribed into RNA, large efforts have been put forth to discover new ways to manipulate this largely untapped area.[2] Many of these approaches have been, or could be used, to generate drugs to manipulate miRNA biology.

2.1 Antisense Oligonucleotides

Antisense Oligonucleotides (ASOs) can be considered the current gold standard for targeting various classes of RNA. ASOs mimic nucleic acids, and utilizing sequence complementary, bind to specific target transcripts. This binding results in changes to transcript levels or their biological processing and function (Figure 2.1).[3] Antisense technology has been used to decrease the expression of a myriad of targets in the research setting for decades due to their relative ease of design and effectiveness;

however, their clinical use has lagged behind.[4] As of 2017, only 4 ASOs have received

FDA approval (Fomivirsen, Mipomersen, Nusinersen, Eteplirsen), all seeing limited use.[5]

One reason the number of FDA-approved ASOs dwarfs that of other classes of drugs is

due to their poor delivery mechanisms and tissue distribution.[6] Antisense oligonucleotides

are relatively large, several thousands in molecular weight, compared to traditional drugs

and contain an extensive amount of negative charges that hinders their cellular uptake.

While various chemistries have been used to improve these properties, most *in vivo* active

ASOs elicit their effects in the blood, liver, or kidneys, or require direct injection into

specific tissues like the spinal cord or retina.[7] This delivery challenge has limited the

diseases for which ASOs are applicable. Additionally, because ASOs mimicking the

structure of DNA and RNA they can be recognized by the immune system triggering
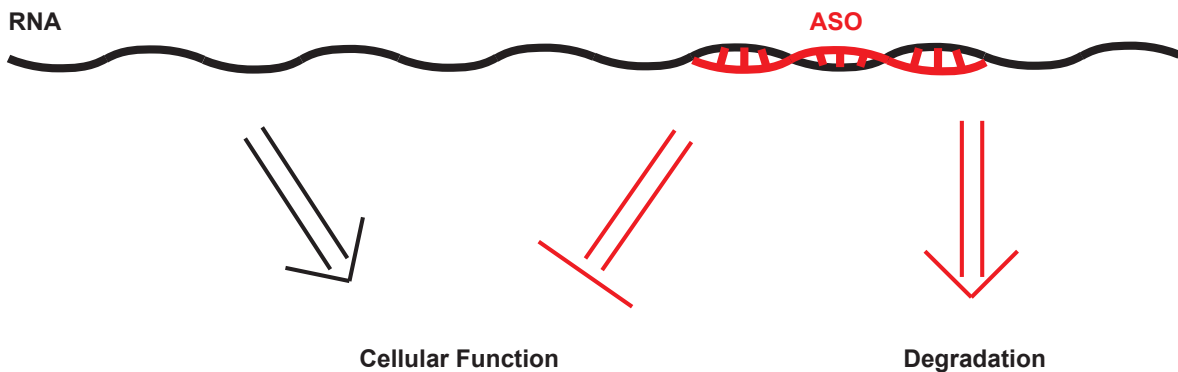
severe immune responses.[8]

Figure 2.1. Antisense Oligonucleotide Schematic. Targeting RNA with antisense oligonucleotides (ASOs)

Despite the challenges with converting ASO technology into FDA-approved

therapies, it has remained the go-to for RNA-based targeting and has been applied to

manipulate miRNA biology. Since the overexpression of miRNAs is common to many

diseases, anti-miRs have been developed to block miRNA function. There are many anti-miRs in development, with two different drug candidates, Miravirsen and RG-101, representing the lead candidates for the first anti-miR approval.[9,10] Both Miravirsen and RG-101 are in phase II trials and target the HCV viral lifecycle through miR-122. These studies have demonstrated that anti-miRs can cause targeted downregulation of their target miRNA humans.[10,11] Alternatively, decreased miRNA expression in various diseases has encouraged the development of miRNA mimetics to supplement these low miRNA levels. One ASO, MesomiR-1, is in phase I as a miR-16 mimic for the treatment of non-small cell lung cancer.[12] Another miRNA mimic for miR-34, MRX34, was also able to reach clinical testing, but was withdrawn due to severe immunological responses.[8] These ASOs provide great proof-of-principle that manipulating miRNAs can have real clinical significance, but the limitations of ASOs have hindered their FDA approval.

2.2 CRISPR

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) is a type of prokaryotic immune system that is being aggressively perused for its broad biomedical applications.[13] CRISPR is famously known for its ability to edit eukaryotic genomes, including humans. This has sparked the development of CRISPR therapies to treat many diseases; however, no human delivery systems have been evaluated.[14] Additionally, there are still major safety concerns with causing permanent changes to the genome. As an alternative to these permanent changes, recent efforts have shown that certain CRISPR systems can selectively target RNA over DNA.[15,16] Building upon this work, Batra et al. demonstrated the potential of a RNA-targeted CRISPR therapy by fusing a nuclease

to the CRIPSR protein Cas9 (Figure 2.2).[17] This nuclease fusion protein lead to transcript degradation of a toxic repeat RNA in cells. Having successfully tested the nuclease degradation of repeat RNAs in cells, the RNA-targeted nuclease system could also be applied to other RNA targets. One example would be to target miRNAs that are overexpressed in diseases by recruiting the nuclease to the miRNA precursors. While many questions still exist about the clinical delivery and off-target effects of CRISPR, there is significant excitement surrounding this new technology.
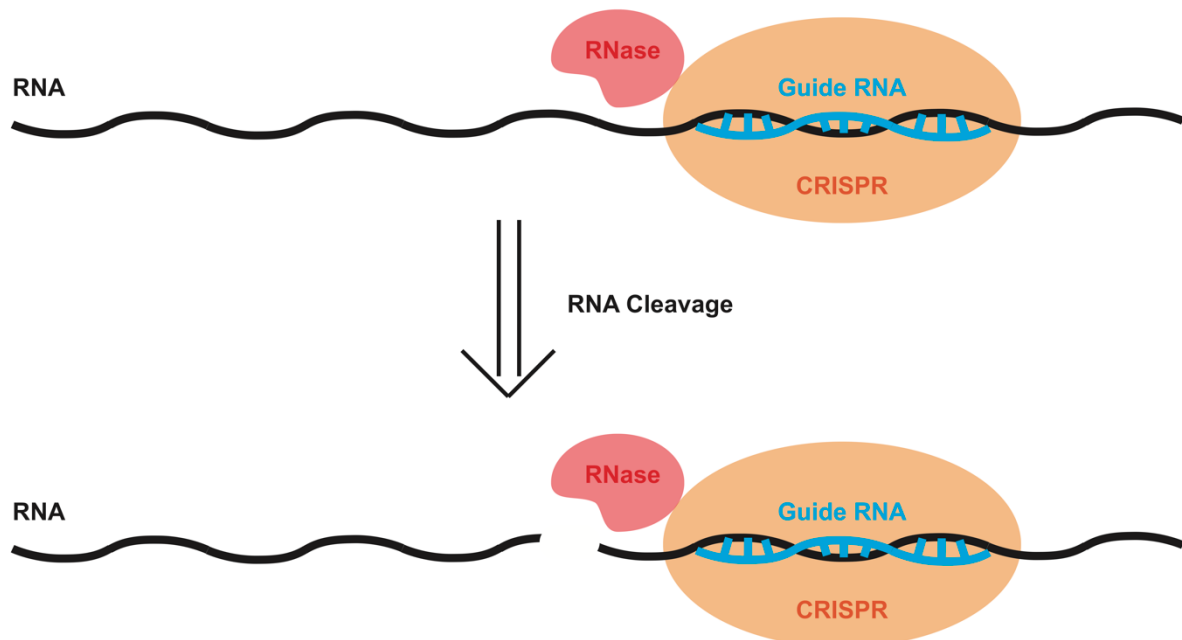

Figure 2.2. Targeting RNA with CRISPR. Targeting RNA with Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) fused to a RNase

2.3 Small Molecules

In contrast to newer technologies like ASOs and CRISPR, the therapeutic use of small molecules was around long before adoption of modern medicine, dating back to the earliest herbal treatments. Small molecules represent 85% of FDA approved drugs largely due to their superior pharmaco-dynamic and -kinetic properties.[1] Interestingly, some of the mainstays of current medicine, such as the tetracyclines and aminoglycosides, function by targeting RNA and have seen widespread use as antibiotics (Figure 2.3). Unfortunately, our ability to target non-ribosomal RNA with small molecules has stalled and fallen behind RNA's growing implications in biology.

streptomycin   Tetracycline
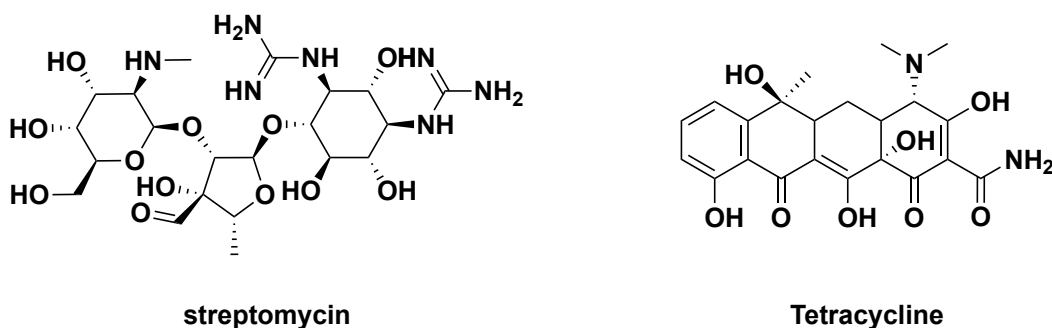
Figure 2.3. Structures of Common RNA Small Molecule Drugs

To date, the largest limitation in using small molecules to target RNA has been obtaining target RNA selectivity.[18,19] This selectivity challenge stems from several properties intrinsic to RNA. First, RNA only consists of four nucleotides, which is in contrast to the 20 amino acids found in proteins. A second challenge is that RNA is fundamentally a polyanionic biomolecule that favors nonspecific electrostatic interactions. Finally, RNAs are known to be highly dynamic, often folding and moving between several

structures, thus making the discovery of traditional druggable spots difficult.[20] However, the identification of high affinity and selective aptamers and riboswitches serve as examples that these challenges can, and have been, overcome.[21,22] To continue to build upon this work and apply this knowledge to disease relevant RNAs, new small molecules, and in particular, chemical scaffolds, need to be discovered. Instrumental to this goal is the development of high-throughput screening (HTS) assays for RNA targets, allowing large numbers of compounds to be rapidly tested. Some of these assays have also been applied to find small molecules regulators of miRNAs and can be categorized into three main groups: cellular, computational, and biochemical.

The generation of endogenous readouts for miRNA activity has enabled the use of cell-based assays for HTS. These reporters utilize a miRNA's ability to suppress gene expression through binding to a 3' untranslated region (UTR) (Figure 2.4A).[23,24] By attaching a readout gene, like luciferase, to a regulatory UTR site, the gene's expression becomes a function of miRNA activity. Compounds that disrupt this activity will derepress the luciferase creating a more intense signal. These cell-based reporters are an excellent way to gauge the functional relevance of test compounds; however, many factors can cause changes in miRNA activity such as changes to transcriptional regulation, miRNA biogenesis, and RISC silencing. This complicates target validation, as it can be difficult to quickly discern the mechanism-of-action for hit compounds. Additionally, cell-based assays can be challenging to adapt to HTS because they require sterile conditions, safety training, and costly media reagents. While there have been compounds successfully identified from these cell based approaches which are summarized in Figure 2.4B, none

have been validated to regulate the RNA directly, and therefore, will not be discussed further as they have been reviewed elsewhere.[19,25-30]
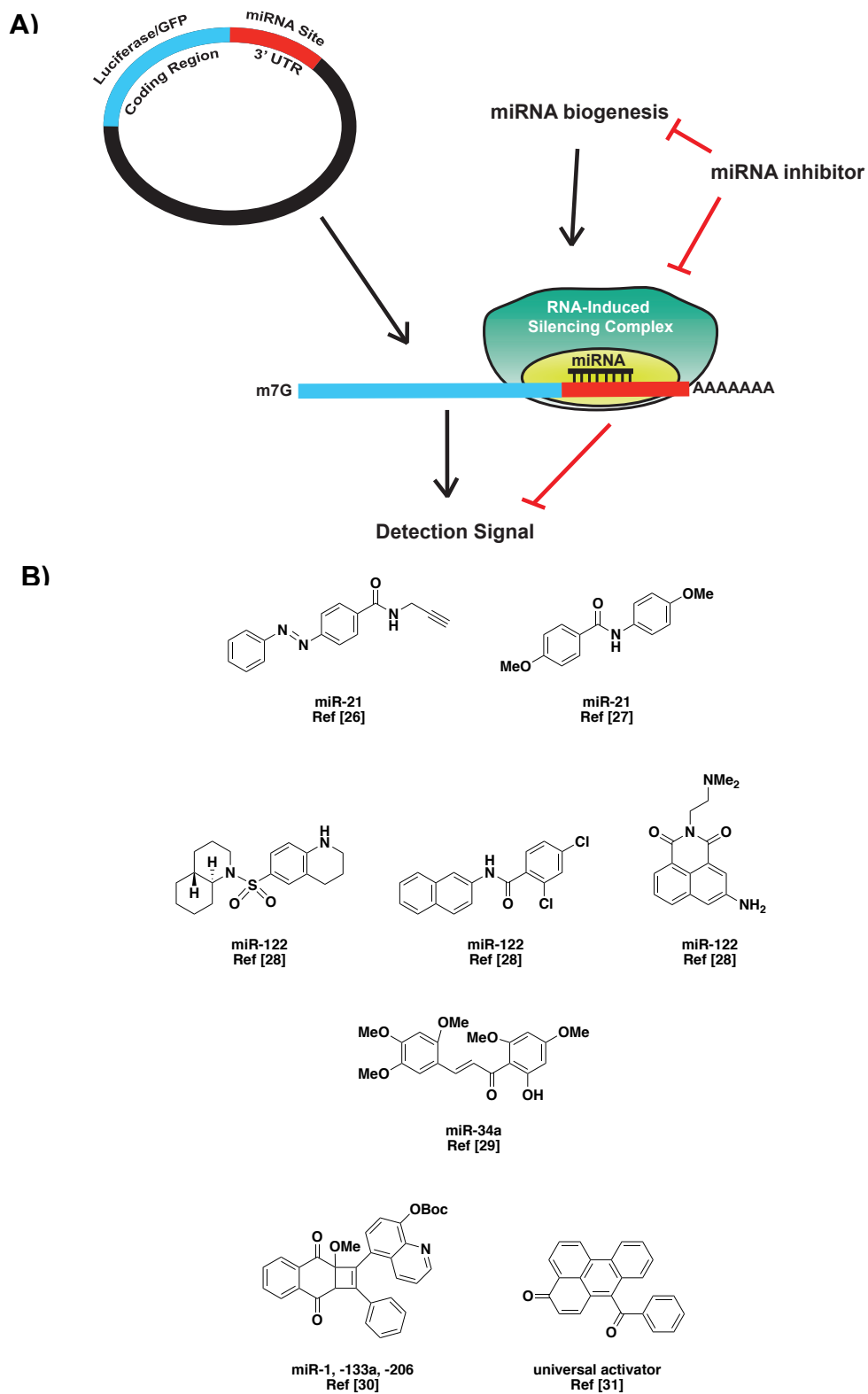
Figure 2.4. Cell Based miRNA Assay and Compounds. A) Depiction of a typical cell-based assay for miRNA activity B) Structures of hits identified from cell based assays

The largest impact on miRNA drug discovery thus far has been from a computational approach pioneered by Professor Mathew Disney and colleagues termed InfoRNA.[31,32] InfoRNA is a database generated from a two-dimensional screen in which small molecules were immobilized in a microarray format and then allowed to bind RNA motifs with internal bulges. Each compound was then analyzed for bound motifs and compiled into a database. After database curation, a target miRNA precursor can be used as an input, analyzed for the presence of internal RNA motifs, outputting small molecules known to bind those sequences. This approach has been used for many different miRNAs, including -18a, -96, -210, -525, and -544, all showing activity in cells and some *in vivo*.[33-37] These experiments have demonstrated that small molecules are capable of targeting and manipulating miRNA biogenesis; however, InfoRNA has its own set of limitations. The current database only consists of 233 small molecules, with a large amount of overlap in chemical space, limiting how many unique scaffolds can be discovered. In fact, this is highlighted in Figure 2.5, as the bis-benzimidazole scaffold was reported in both the inhibition miR-96 and -210 by targeting different internal motifs, CGA and ACU, respectively.[34,35] This has also brought RNA selectivity back into question. Expansion of the InfoRNA library is challenging because it requires the compounds to be immobilized which requires chemical handles not present in large scale compound libraries.
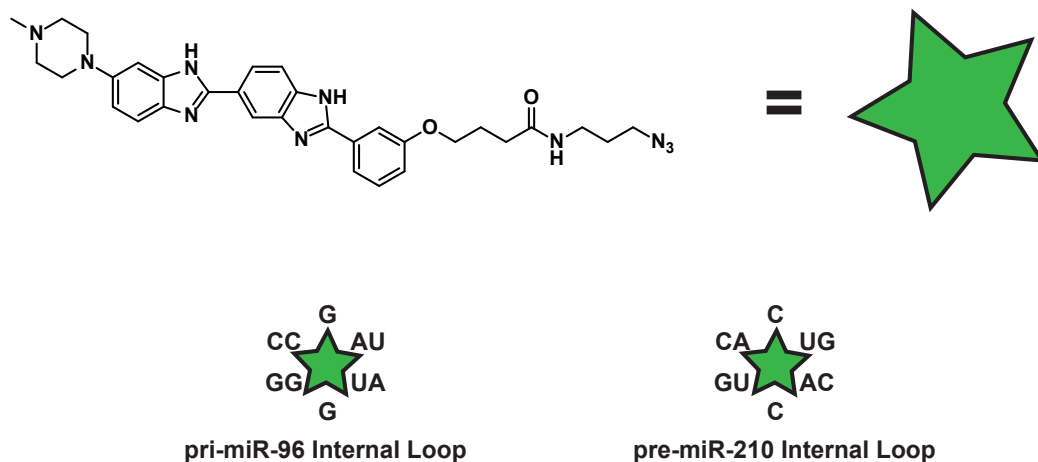
Figure 2.5. Comparing Bisbenzimidazole RNA Targets. Bisbenzimidazole used by the Disney group to target two different internal RNA loops

In contrast to the previous two strategies, standard biochemical approaches use traditional compound libraries to assay a target directly. Many different biochemical assays have been developed for miRNAs, including FRET, microarray, and binding or displacement assays. All of these techniques have been reviewed in detail elsewhere, and some are represented in Figure 2.6.[19] While successfully implemented, these assays have yet to be tested in large scale screening campaigns >50,000 compounds. One problem associated with scaling-up these assays is that their fluorescent-based detection method is prone to compound interference from naturally fluorescent or fluorescent quenching molecules. Additionally, the majority of these assays do not assess function, meaning hit compounds will not necessarily result in the desired functional activity. While these drawbacks can be overcome by further analysis of hits, it would require a significant amount of effort to filter the thousands of hits commonly found in a large HTS campaign. Similar to InfoRNA, previous biochemical screens for miRNAs have focused on existing

chemical space, consisting of aminoglycosides and intercalators. The problem with testing aminoglycosides and their derivatives is that their positive charge makes selectivity for one RNA over another challenging, as binding is largely dependent on electrostatic interactions. As shown by Figure 2.7, there are many reported miRNA modulators that rely on positive charge and likely exhibit large off-target effects.[36,38-45] Intercalators also suffer from lack of selectivity, as their binding energies are derived from pi-stacking inside the RNA or DNA helix. The continued use of these molecules, despite their nonspecific nature, emphasizes the need for new types of RNA scaffolds and mechanisms to be uncovered to target these important biomolecules.
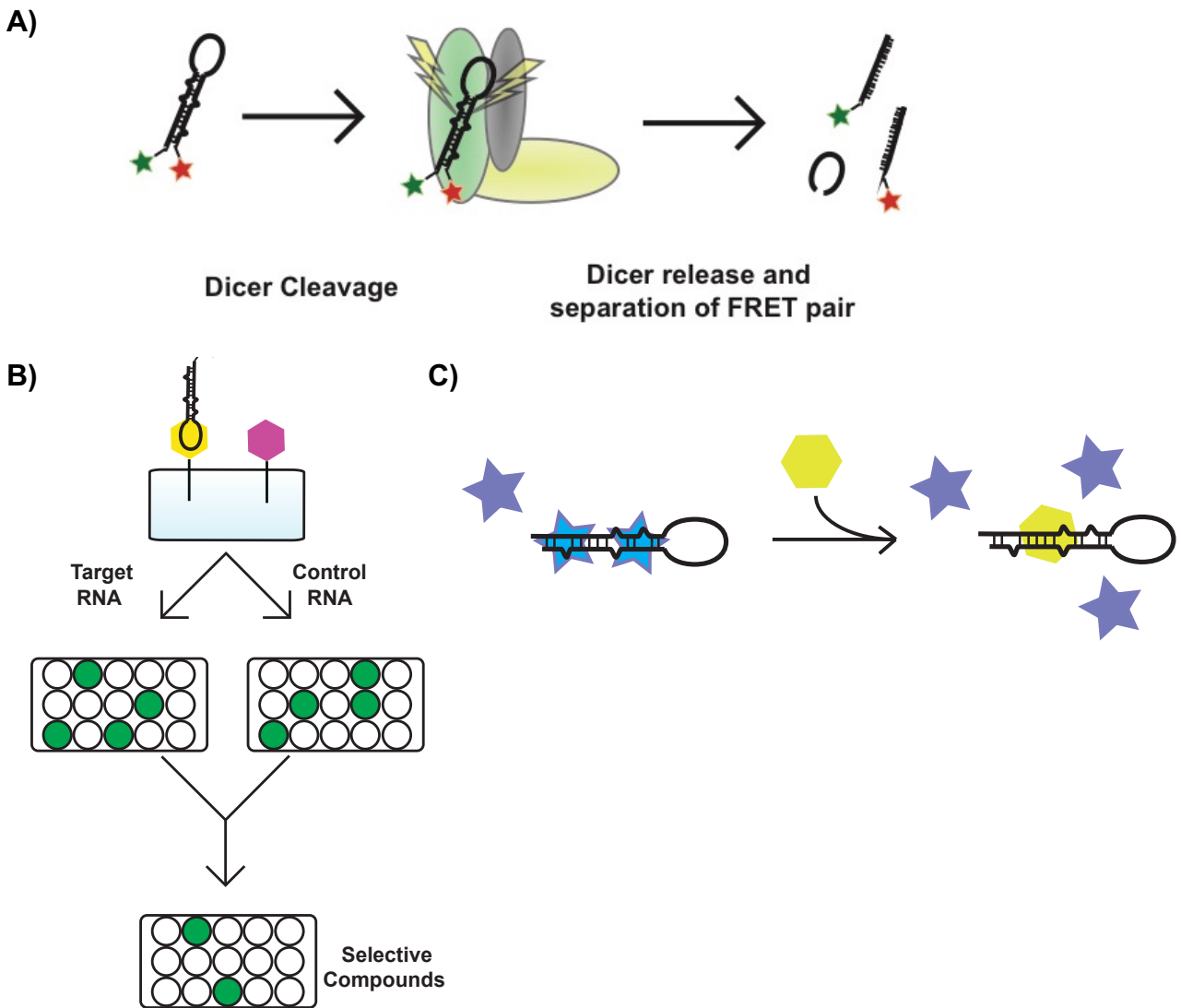
Figure 2.6. In Vitro miRNA Assays. A) FRET based detection B) Microarray C) Fluorescent displacement assay

pri-miR-21
Ref [39]

pri-miR-21
Ref [40]

pri-miR-155
Ref [41]

miR-10b
Ref [37]

miR-21
Ref [42]

miR-525
Ref [43]

pre-let-7
Ref [44]

pre-miR-372/373
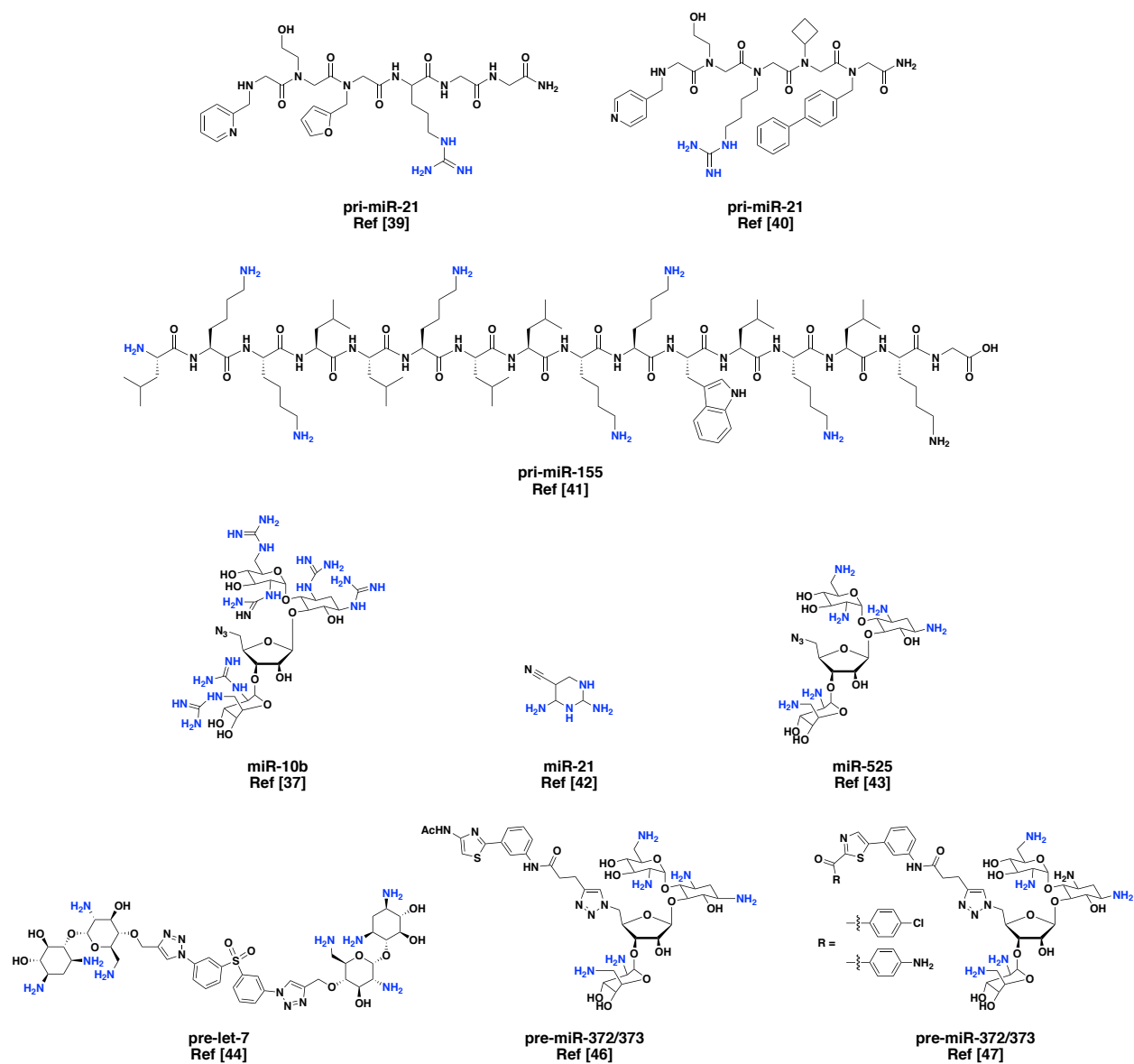**Ref [46]**

pre-miR-372/373
Ref [47]

Figure 2.7. Positively Charged miRNA Inhibitors.

2.4 Small Molecule Targeting of RNA Binding Proteins

As an alternative to targeting RNAs directly, one up and coming strategy is to utilize the specificity already found in endogenous RNA-Binding Proteins (RBPs). It has been estimated that 1,542 proteins contain RNA-binding motifs, many of which have been characterized with defined binding sequences.[46] By targeting these RBPs with small molecules, one could affect the biology of their RNA targets. While this strategy has been demonstrated for other RBP systems like splicing and viral infection, the RBP Lin28 will be focused on for its role in regulating the biogenesis of the miRNA let-7.[47,48]

To date there have been three reported screens targeting the Lin28-let-7 interaction. Roos et al. published the first screen based on a FRET reporter between a Lin28b-green fluorescent protein (GFP) fusion protein and a Black-Hole-Quencher 1 (BHQ-1) labeled truncated pre-let-7.[49] Using this FRET system, they screened a commercial library of 16,000 compounds and were able to identify one molecule, **1**, with single digit micromolar activity *in vitro*. Confirming the viability of targeting a RBP, the compound was able to increase let-7 levels in cells by inhibiting Lin28-let-7 resulting in decreased Lin28 expression. Shortly after this, Lim et al. used a similar FRET assay, but they incorporated an unnatural amino acid for site specific labeling instead of generating a fusion protein.[50] After screening 4,500 compounds, the authors were able to find a molecule, **2**, that binds the cold shock domain of Lin28 with single digit nanomolar affinity which was active *in celluo* at 40 μM. Finally, Lightfoot et al. generated a fluorescence polarization assay to screen 2,768 compounds against Lin28-let-7.[51] From this screen

they were able to identify several scaffolds, **3-5**, with *in vitro* activity, but the cellular activity of these molecules was not evaluated.
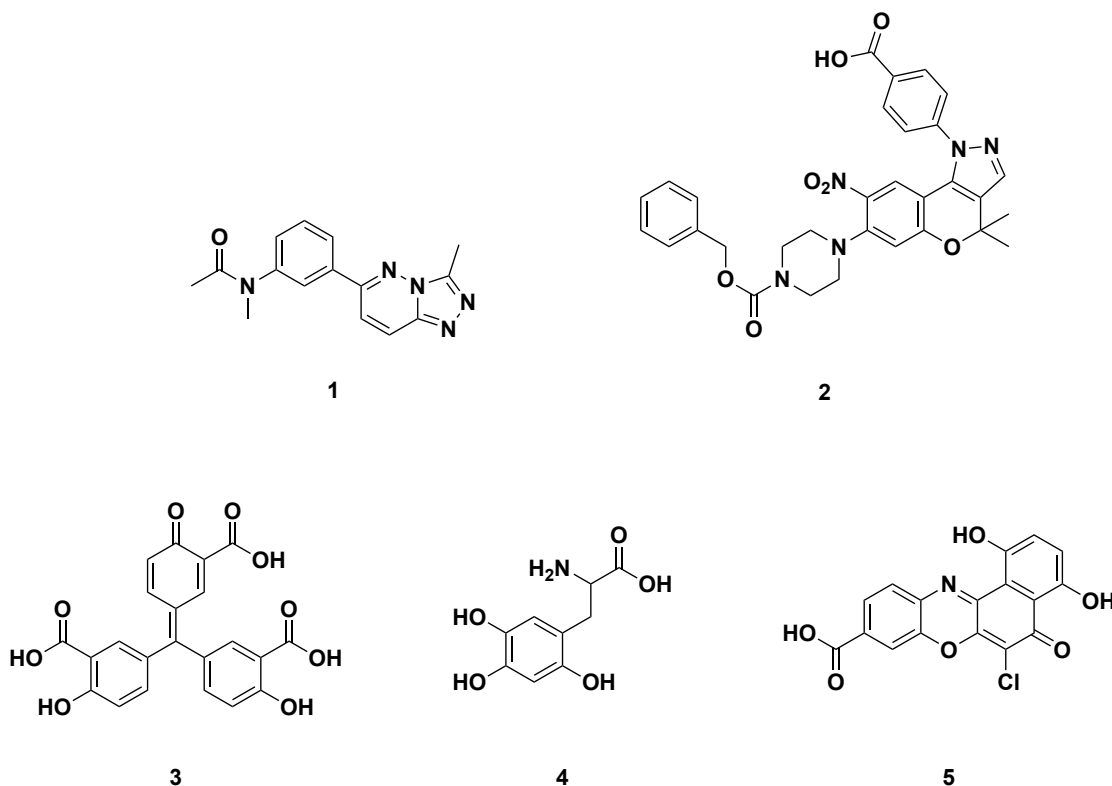
Figure 2.8. Structures of pre-let7-Lin28 Inhibitors.

2.5 Concluding Remarks

The therapeutic benefit of targeting RNA, and in particular miRNAs, continues to expand rapidly. While current technologies have demonstrated that RNAs are in fact valid clinical targets, our ability to manipulate RNA has fallen short. New approaches like ASOs and CRISPR show great promise, but have already shown clinical limitations, and the traditional small molecule strategy has suffered from lack of scalability and desire to screen new chemical space. Taken together, this highlights the need for new drug discovery technologies to be developed to take advantage of our growing understanding of RNA.

2.6 Copyright

The work in this chapter was reproduced in part from Lorenz, D. A. & Garner, A. L. in *RNA Therapeutics* (ed. Garner, A. L.) 79–110 (Springer International Publishing, 2018).

2.7 References

1.   Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I. & Overington, J. P. A comprehensive map of molecular drug targets. *Nature reviews. Drug discovery* **16,** 19–34 (2017).

2.   Connelly, C. M., Moon, M. H. & Schneekloth, J. S. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chemical Biology* **23,** 1077–1090 (2016).

3.   Li, Z. & Rana, T. M. Therapeutic targeting of microRNAs: current status and future challenges. *Nature Publishing Group* **13,** 622–638 (2014).

4.   Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391,** 806–811 (1998).

5.   Stein, C. A. & Castanotto, D. FDA-Approved Oligonucleotide Therapies in 2017. *Mol. Ther.* **25,** 1069–1075 (2017).

6.   Geary, R. S., Norris, D., Yu, R. & Bennett, C. F. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Advanced Drug Delivery Reviews* **87,** 46–51 (2015).

7.   Wan, W. B. & Seth, P. P. The Medicinal Chemistry of Therapeutic Oligonucleotides. *J. Med. Chem.* **59,** 9645–9667 (2016).

8.   Chakraborty, C., Sharma, A. R., Sharma, G., Doss, C. G. P. & Lee, S.-S. Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine. *Mol Ther Nucleic Acids* **8,** 132–143 (2017).

9.   van der Ree, M. H., van der Meer, A. J., de Bruijne, J., Maan, R., van Vliet, A., Welzel, T. M., Zeuzem, S., Lawitz, E. J., Rodriguez-Torres, M., Kupcova, V., Wiercinska-Drapalo, A., Hodges, M. R., Janssen, H. L. A. & Reesink, H. W. Long-

term safety and efficacy of microRNA-targeted therapy in chronic hepatitis C patients. *Antiviral Res.* **111,** 53–59 (2014).

10. van der Ree, M. H., de Vree, J. M., Stelma, F., Willemse, S., van der Valk, M., Rietdijk, S., Molenkamp, R., Schinkel, J., van Nuenen, A. C., Beuers, U., Hadi, S., Harbers, M., van der Veer, E., Liu, K., Grundy, J., Patick, A. K., Pavlicek, A., Blem, J., Huang, M., Grint, P., Neben, S., Gibson, N. W., Kootstra, N. A. & Reesink, H. W. Safety, tolerability, and antiviral effect of RG-101 in patients with chronic hepatitis C: a phase 1B, double-blind, randomised controlled trial. *Lancet* **389,** 709–717 (2017).

11. van der Ree, M. H., van der Meer, A. J., van Nuenen, A. C., de Bruijne, J., Ottosen, S., Janssen, H. L., Kootstra, N. A. & Reesink, H. W. Miravirsen dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Aliment. Pharmacol. Ther.* **43,** 102–113 (2016).

12. Reid, G., Kao, S. C., Pavlakis, N., Brahmbhatt, H., MacDiarmid, J., Clarke, S., Boyer, M. & van Zandwijk, N. Clinical development of TargomiRs, a miRNA mimic-based treatment for patients with recurrent thoracic cancer. *Epigenomics* **8,** 1079–1085 (2016).

13. Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annual review of biochemistry* **85,** 227–264 (2016).

14. Cai, L., Fisher, A. L., Huang, H. & Xie, Z. CRISPR-mediated genome editing and human diseases. *Genes & Diseases* **3,** 244–251 (2016).

15. O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M. & Doudna, J. A. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516,** 263–266 (2014).

16. Nelles, D. A., Fang, M. Y., O'Connell, M. R., Xu, J. L., Markmiller, S. J., Doudna, J. A. & Yeo, G. W. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **165,** 488–496 (2016).

17. Batra, R., Nelles, D. A., Pirie, E., Blue, S. M., Marina, R. J., Wang, H., Chaim, I. A., Thomas, J. D., Zhang, N., Nguyen, V., Aigner, S., Markmiller, S., Xia, G., Corbett, K. D., Swanson, M. S. & Yeo, G. W. Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell* **170,** 899–912.e10 (2017).

18. Matsui, M. & Corey, D. R. Non-coding RNAs as drug targets. *Nature Publishing Group* **16,** 167–179 (2017).

19. Lorenz, D. A. & Garner, A. L. in *RNA Therapeutics* (ed. Garner, A. L.) 79–110 (Springer International Publishing, 2018).

20. Getz, M., Sun, X., Casiano-Negroni, A., Zhang, Q. & Al-Hashimi, H. M. NMR studies of RNA dynamics and structural plasticity using NMR residual dipolar couplings. *Biopolymers* **86,** 384–402 (2007).

21. Werstuck, G. & Green, M. R. Controlling Gene Expression in Living Cells Through Small Molecule-RNA Interactions. *Science* **282,** 296–298 (1998).

22. Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441,** 1167–1171 (2006).

23. Connelly, C. M., Thomas, M. & Deiters, A. High-throughput luciferase reporter assay for small-molecule inhibitors of microRNA function. *Journal of Biomolecular Screening* **17,** 822–828 (2012).

24. Connelly, C. M. & Deiters, A. in *miRNA Maturation* (ed. Arenz, C.) **1095,** 147–156 (Humana Press, 2013).

25. Gumireddy, K., Young, D. D., Xiong, X., Hogenesch, J. B., Huang, Q. & Deiters, A. Small-molecule inhibitors of microrna miR-21 function. *Angewandte Chemie (International ed. in English)* **47,** 7482–7484 (2008).

26. Naro, Y., Thomas, M., Stephens, M. D., Connelly, C. M. & Deiters, A. Aryl amide small-molecule inhibitors of microRNA miR-21 function. *Bioorganic & Medicinal Chemistry Letters* **25,** 4793–4796 (2015).

27. Small molecule modifiers of microRNA miR-122 function for the treatment of hepatitis C virus infection and hepatocellular carcinoma. **132,** 7976–7981 (2010).

28. Xiao, Z., Li, C. H., Chan, S. L., Xu, F., Feng, L., Wang, Y., Jiang, J.-D., Sung, J. J. Y., Cheng, C. H. K. & Chen, Y. A small-molecule modulator of the tumor-suppressor miR34a inhibits the growth of hepatocellular carcinoma. *Cancer Research* **74,** 6236–6247 (2014).

29. Tan, S.-B., Huang, C., Chen, X., Wu, Y., Zhou, M., Zhang, C. & Zhang, Y. Small molecular inhibitors of miR-1 identified from photocycloadducts of acetylenes with 2-methoxy-1,4-naphthalenequinone. *Bioorganic & Medicinal Chemistry* **21,** 6124–6131 (2013).

30. Chen, X., Huang, C., Zhang, W., Wu, Y., Chen, X., Zhang, C.-Y. & Zhang, Y. A universal activator of microRNAs identified from photoreaction products. *Chem. Commun.* **48,** 6432–6434 (2012).

31. Velagapudi, S. P., Gallo, S. M. & Disney, M. D. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nature Chemical*

Biology **10,** 291–297 (2014).

32.   Disney, M. D., Winkelsas, A. M., Velagapudi, S. P., Southern, M., Fallahi, M. & Childs-Disney, J. L. Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules Targeting Structured RNAs. *ACS chemical biology* **11,** 1720–1728 (2016).

33.   Velagapudi, S. P., Luo, Y., Tran, T., Haniff, H. S., Nakai, Y., Fallahi, M., Martinez, G. J., Childs-Disney, J. L. & Disney, M. D. Defining RNA-Small Molecule Affinity Landscapes Enables Design of a Small Molecule Inhibitor of an Oncogenic Noncoding RNA. *ACS Cent Sci* **3,** 205–216 (2017).

34.   Velagapudi, S. P., Cameron, M. D., Haga, C. L., Rosenberg, L. H., Lafitte, M., Duckett, D. R., Phinney, D. G. & Disney, M. D. Design of a small molecule against an oncogenic noncoding RNA. *Proceedings of the National Academy of Sciences* **113,** 5898–5903 (2016).

35.   Costales, M. G., Haga, C. L., Velagapudi, S. P., Childs-Disney, J. L., Phinney, D. G. & Disney, M. D. Small Molecule Inhibition of microRNA-210 Reprograms an Oncogenic Hypoxic Circuit. *Journal of the American Chemical Society* jacs.6b11273–10 (2017). doi:10.1021/jacs.6b11273

36.   Childs-Disney, J. L. & Disney, M. D. Small Molecule Targeting of a MicroRNA Associated with Hepatocellular Carcinoma. *ACS chemical biology* **11,** 375–380 (2016).

37.   Haga, C. L., Velagapudi, S. P., Strivelli, J. R., Yang, W.-Y., Disney, M. D. & Phinney, D. G. Small Molecule Inhibition of miR-544 Biogenesis Disrupts Adaptive Responses to Hypoxia by Modulating ATM-mTOR Signaling. *ACS chemical biology* **10,** 2267–2276 (2015).

38.   Chirayil, S., Chirayil, R. & Luebke, K. J. Discovering ligands for a microRNA precursor with peptoid microarrays. *Nucleic Acids Research* **37,** 5486–5497 (2009).

39.   Diaz, J. P., Chirayil, R., Chirayil, S., Tom, M., Head, K. J. & Luebke, K. J. Association of a peptoid ligand with the apical loop of pri-miR-21 inhibits cleavage by Drosha. *RNA* **20,** 528–539 (2014).

40.   Pai, J., Hyun, S., Hyun, J. Y., Park, S.-H., Kim, W.-J., Bae, S.-H., Kim, N.-K., Yu, J. & Shin, I. Screening of Pre-miRNA-155 Binding Peptides for Apoptosis Inducing Activity Using Peptide Microarrays. *Journal of the American Chemical Society* **138,** 857–867 (2016).

41.   Velagapudi, S. P. & Disney, M. D. Two-dimensional combinatorial screening enables the bottom-up design of a microRNA-10b inhibitor. *Chem. Commun.* **50,**

3027–3029 (2014).

42.   Shi, Z., Zhang, J., Qian, X., Han, L., Zhang, K., Chen, L., Liu, J., Ren, Y., Yang, M., Zhang, A., Pu, P. & Kang, C. AC1MMYR2, an inhibitor of dicer-mediated biogenesis of Oncomir miR-21, reverses epithelial-mesenchymal transition and suppresses tumor growth and progression. *Cancer Research* **73,** 5519–5531 (2013).

43.   Klemm, C. M., Berthelmann, A., Neubacher, S. & Arenz, C. Short and Efficient Synthesis of Alkyne-Modified Amino Glycoside Building Blocks. *Eur. J. Org. Chem.* **2009,** 2788–2794 (2009).

44.   Vo, D. D., Staedel, C., Zehnacker, L., Benhida, R., Darfeuille, F. & Duca, M. Targeting the production of oncogenic microRNAs with multimodal synthetic small molecules. *ACS chemical biology* **9,** 711–721 (2014).

45.   Vo, D. D., Tran, T. P. A., Staedel, C., Benhida, R., Darfeuille, F., Di Giorgio, A. & Duca, M. Oncogenic MicroRNAs Biogenesis as a Drug Target: Structure-Activity Relationship Studies on New Aminoglycoside Conjugates. *Chemistry* **22,** 5350–5362 (2016).

46.   Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nature Publishing Group* **15,** 829–845 (2014).

47.   Palacino, J., Swalley, S. E., Song, C., Cheung, A. K., Shu, L., Zhang, X., Van Hoosear, M., Shin, Y., Chin, D. N., Keller, C. G., Beibel, M., Renaud, N. A., Smith, T. M., Salcius, M., Shi, X., Hild, M., Servais, R., Jain, M., Deng, L., Bullock, C., McLellan, M., Schuierer, S., Murphy, L., Blommers, M. J. J., Blaustein, C., Berenshteyn, F., Lacoste, A., Thomas, J. R., Roma, G., Michaud, G. A., Tseng, B. S., Porter, J. A., Myer, V. E., Tallarico, J. A., Hamann, L. G., Curtis, D., Fishman, M. C., Dietrich, W. F., Dales, N. A. & Sivasankaran, R. SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nature Chemical Biology* **11,** 511–517 (2015).

48.   Mei, H. Y., Mack, D. P., Galan, A. A., Halim, N. S., Heldsinger, A., Loo, J. A., Moreland, D. W., Sannes-Lowery, K. A., Sharmeen, L., Truong, H. N. & Czarnik, A. W. Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorganic & Medicinal Chemistry* **5,** 1173–1184 (1997).

49.   Roos, M., Pradère, 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 chemical biology* **11,** 2773–2781 (2016).

50.   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* **138,** 13630–13638 (2016).

51.   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.* **14,** 10208–10216 (2016).

Chapter 3

Small Molecule Targeting of pre-miRNAs

As discussed in Chapter 1, the regulation of miRNA biogenesis is critical for maintaining healthy physiology. Due to miRNA's role in diseases, substantial efforts have been put forth to find ways to manipulate miRNA expression for therapeutic benefit. Unfortunately, as highlighted in Chapter 2, current small molecule RNA probes tend to lack sufficient selectivity. This can, in part, be attributed to the repurposing of known RNA scaffolds instead of identifying new RNA-targeted chemical space. However, the discovery of new small molecules is also complicated by inadequate assays to quickly preform reliable high-throughput screens (HTS). Therefore, to discover new small molecules and scaffolds for the manipulation of Dicer-dependent miRNA maturation, a new, sensitive, readily adaptable, and HTS-compatible assay was developed.

Catalytic Enzyme-Linked Click Chemistry Assay, or cat-ELCCA, is a transformative technology that utilizes click chemistry to overcome many of the challenges associated with traditional HTS-compatible assays.[1-6] In general, cat-ELCCA is similar to a standard ELISA (enzyme-linked immunosorbent assay), but does not require antibodies with a generic scheme shown in Figure 3.1.[7] Click chemistry is uniquely suited to take the place of costly and potentially challenging-to-generate antibodies because click reactions are defined by their high yields, physiological stability, minimal

and inoffensive by-products, readily available reagents, and minimal solvent requirement.[8] These ideal reaction properties allow for selective and covalent attachment of an enzyme to an analyte-of-interest, resulting in robust catalytic signal amplification. Additionally, by immobilizing the analyte, compounds that are known to interfere with other assays are removed by a simple washing step, thus reducing follow-up time. Furthermore, the assay designs are constructed to report functional activity as opposed to many assays only reporting simple binding events. Finally, cat-ELCCA is compatible with traditional HTS platforms, allowing for screening of much larger and diverse libraries when compared to other assay formats such as microarrays. These benefits were highlighted by the first cat-ELCCA. Conceived for the monitoring of Ghrelin O-Acyltransferase (GOAT) activity, this cat-ELCCA used copper-catalyzed click chemistry to attach a functionalized HRP to a peptide fragment and allowed Drs. Amanda Garner and Kim Janda to identify the first small molecule inhibitor of GOAT.[1,2]
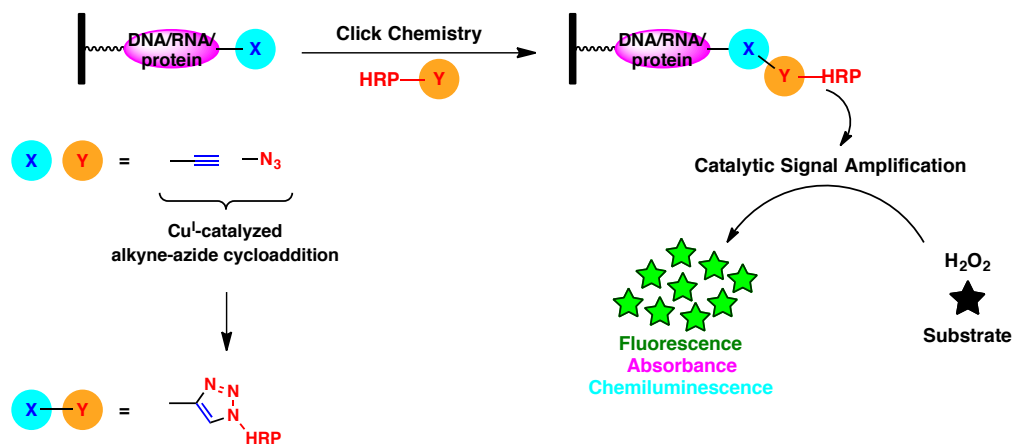

Figure 3.1. General schematic for a cat-ELCCA assay.

3.1 Dicer cat-ELCCA

Inspired by the successful implementation of cat-ELCCA for GOAT, a new cat-ELCCA was designed and developed for Dicer-dependent maturation of pre-miRNAs (Figure 3.2A).[3] To briefly describe the design, a pre-miRNA hairpin loop containing a click handle is first immobilized in a HTS-compatible microtiter plate. Compounds are then added prior to the addition of Dicer to allow compounds to bind the RNA before being subjected to Dicer cleavage. After a washing step to remove digested RNA, compounds, and Dicer, a derivatized Horseradish Peroxidase (HRP) is then added to react with uncleaved pre-miRNAs, generating a RNA-HRP conjugate through click chemistry. Before the addition of the HRP substrate, unreacted HRP is removed by washing. This step makes the HRP activity directly correlated with the amount of uncleaved RNA and therefore Dicer activity. This new assay brings all of the benefits discussed earlier and enables quick adaptation between pre-miRNA targets as the RNA can be easily changed without changing the procedure. The design of the RNA substrate only requires two modifications, both of which are incorporated during chemical synthesis and are commercially available. The first modification is a biotin tag with a polyethylene glycol linker that functions as the immobilization handle and spacer from the surface. This linker location was chosen at the 5' end of the sequence to avoid interfering with Dicer recognition of the 2 nucleotide 3' overhang; although, the 3' end modification was not tested to confirm this locational requirement.[9] The second modification is an aminoallyl uridine located in the terminal loop region that can be derivatized by N-Hydroxysuccinimide (NHS) coupling to incorporate a click chemistry handle, allowing for detection of Dicer cleavage (Figure 3.2B).
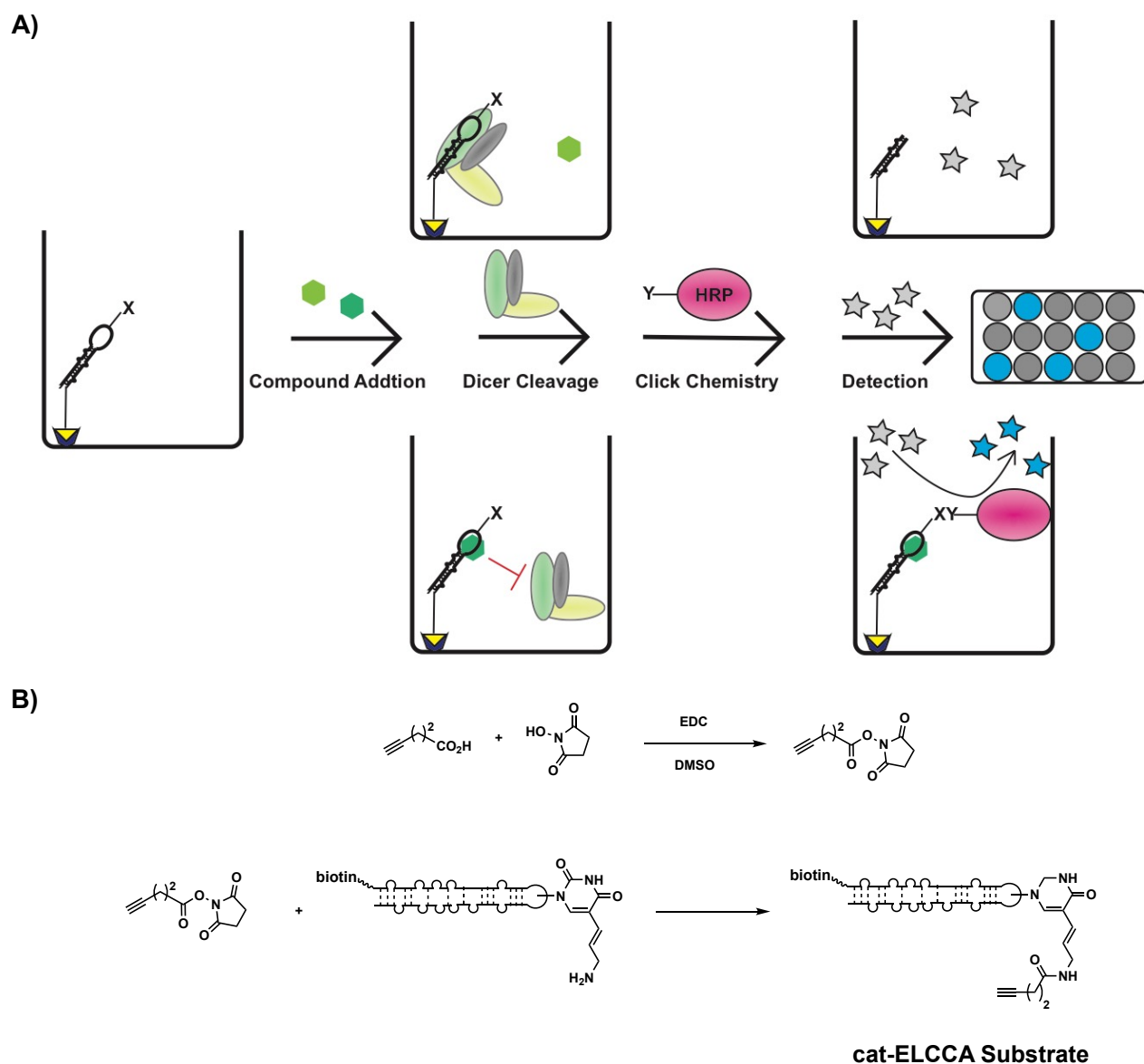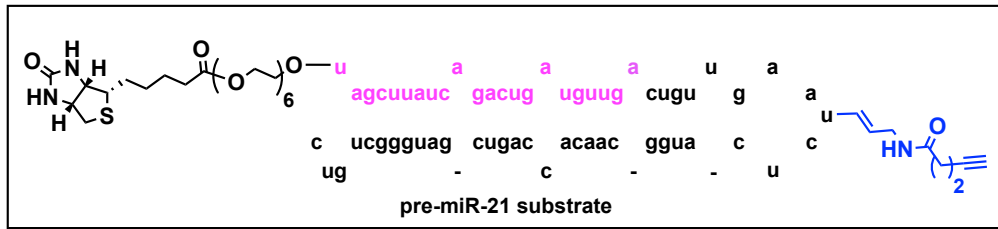
Figure 3.2. Dicer cat-ELCCA. A) Dicer cat-ELCCA scheme B) Alkyne coupling to the modified

This new cat-ELCCA design was then applied to the discovery of biogenesis inhibitors for miR-21. Chosen for its role in cancer, this overexpressed miRNA serves a prime candidate for small molecule-based inhibitors.[10] To accomplish this, several components were synthesized and tested to ensure the Dicer cat-ELCCA would function properly. As previously discussed, the first step was to conjugate the RNA substrate to

an alkyne click chemistry handle through standard NHS chemistry to generate a cat-ELCCA ready substrate (Figure 3.3A). This cat-ELCCA-ready RNA substrate was then reacted with a rhodamine dye functionalized with an azide to generate a fluorophore-RNA conjugate, which was confirmed by fluorescent gel imaging. To make sure that the modified cat-ELCCA substrate was still able to undergo Dicer cleavage, the substrate was incubated with a commercially sourced Dicer. Analysis by denaturing gel electrophoresis revealed Dicer was tolerant of these modifications by cleaving the substrate to mature miRNAs (Figure 3.3B). To prove that the miRNA formation was due to Dicer activity, Dicer was inactivated by boiling and the addition of a metal chelator to remove the catalytic magnesium ions. As expected, no enzymatic activity was observed with inactivated Dicer. Next, the immobilization efficiency of the substrate was evaluated by adding the RNA into a streptavidin-coated microtiter plate, incubated overnight to immobilize, and then the supernatant was removed. The non-immobilized RNA in the supernatant was then quantified by gel electrophoresis (Figure 3.3C and D). This quantification revealed that the cat-ELCCA substrate underwent successful immobilization; although, less than half the theoretical capacity was occupied. The final component for cat-ELCCA, an azide-functionalized HRP, was generated based on a previously reported method and the presence of the click handle was confirmed by conjugation to a rhodamine click chemistry derivative (Figure 3.3E and F).
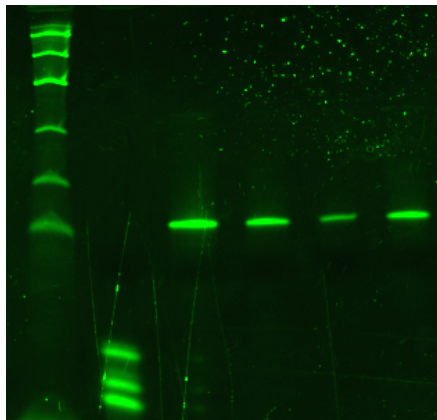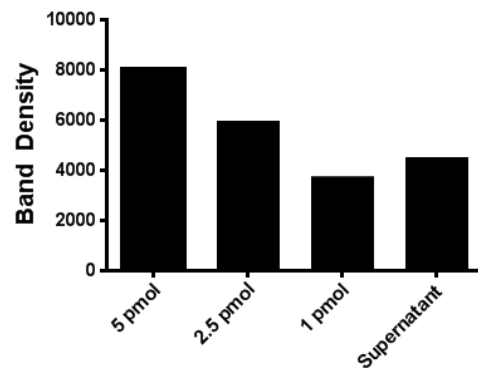
A)

pre-miR-21 substrate
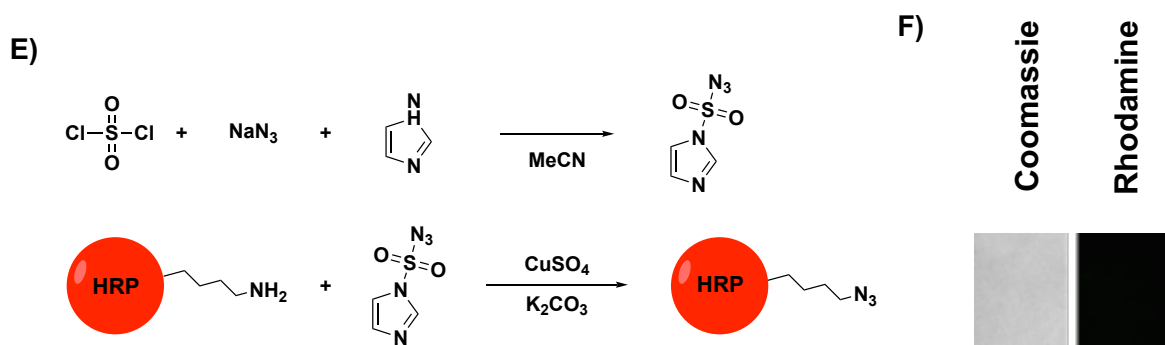
B)
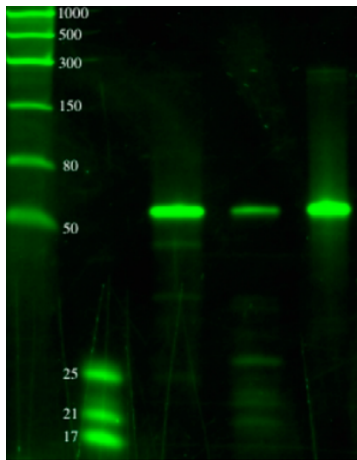
C)

D)

Figure 3.3. Preliminary Dicer cat-ELCCA Experiments. A) pre-miR-21-alkyne cat-ELCCA substrate B) In Vitro digest of pre-miR-21 substrate with commercial Dicer. Gel lanes from left to right are: Low MW ssRNA ladder, miRNA ladder, inactivated Dicer, Dicer, and no Dicer. C) Quantification of remaining RNA substrate after overnight immobilization. Gel lanes from left to right are: Low MW ssRNA ladder, miRNA ladder, 5 pmol substrate standard, 2.5 pmol substrate standard, 1 pmol substrate standard, remaining substrate. D) Quantification by ImageJ of gel bands from C. E) Synthesis of azide functionalized HRP. F) Confirmation of azide functionalized HRP by reaction with a rhodamine alkyne and subsequent coomassie and rhodamine visualization.

With all of the components necessary for cat-ELCCA synthesized and tested for activity, several proof-of-concept experiments were performed. Following the design outlined previously, the appropriate amount of enzyme and time required for optimal signal was first determined. As seen in Figure 3.4A, 0.25 units at 6 hours produced the best results. Importantly, this change in signal was lost when inactivated Dicer was used, matching results seen in solution (Figure 3.4B). It should be noted that this is the first time a cat-ELCCA format was used in a 384-well plate, a standard format for HTS. Additionally, improvements were seen by switching to a chemiluminescence detection method from the previous fluorescent-based detection. To demonstrate the benefits of the cat-ELCCA platform over traditional assays, dichlorofluorescein and guanine diphosphate were tested. These compounds are known to interfere with most assays due to their fluorescent

and fluorescent quenching properties respectively.[11] As expected, neither of the common interfering mechanisms affected the assay with either a fluorescent or chemiluminescent readout (Figure 3.4C). Unexpectedly, dichlorofluorescein had low levels of inhibitory activity. This example demonstrates that cat-ELCCA can be used to discover molecules that would likely be excluded from other assays based on their optical properties, not on their functional activity. Finally, to assess the assay's reproducibility, 48 wells of no Dicer and Dicer controls were tested, which resulted in Z' value of 0.6. Z' values are commonly used to evaluate an assay's suitability for HTS, as it represents both signal-to-background and standard deviations, with Z' values greater than 0.5 considered excellent assays for HTS.[12] The combination of these experiments validated that cat-ELCCA can serve as a reliable platform for detecting Dicer-dependent miRNA maturation.
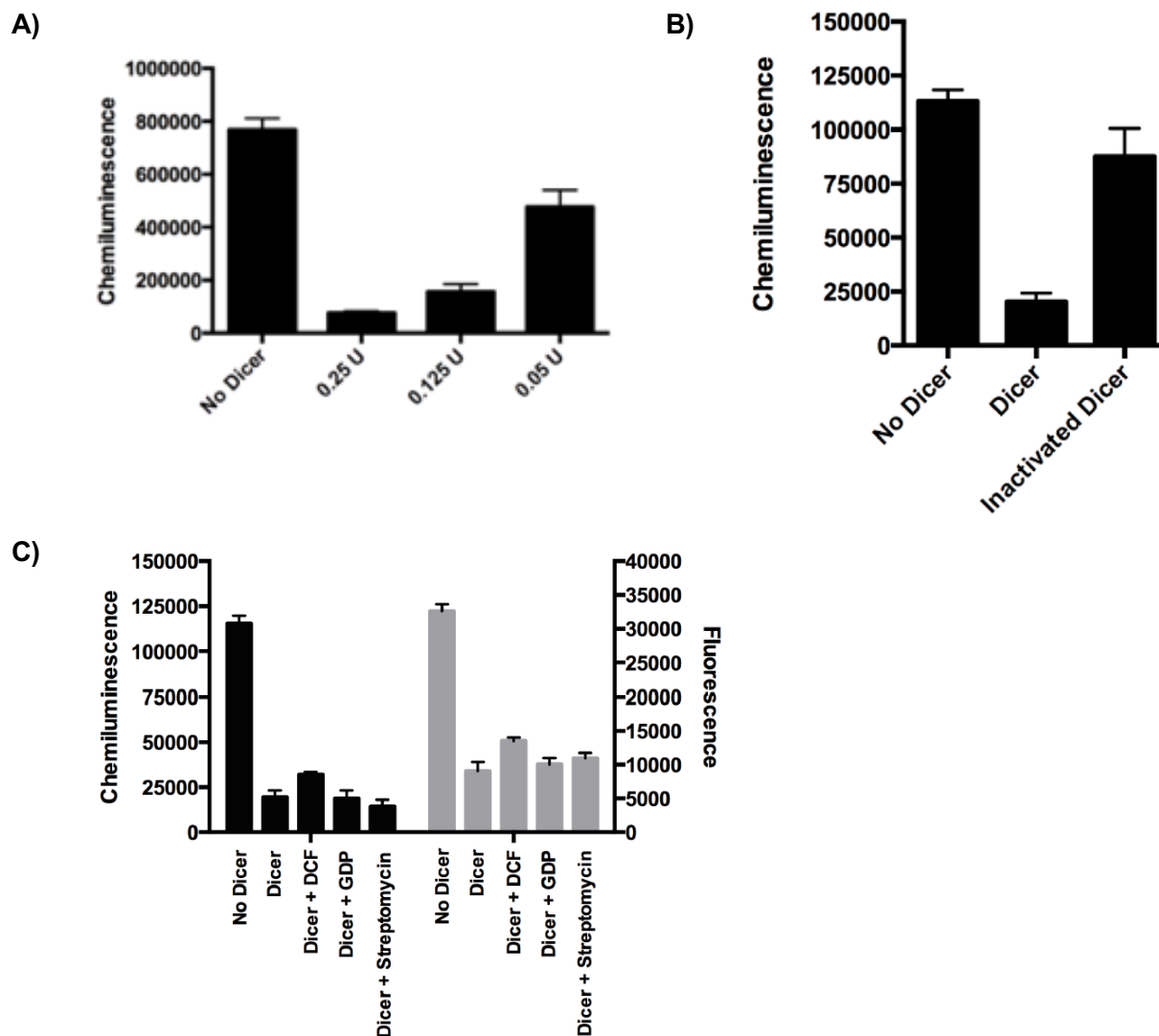
Figure 3.4. Proof-of-Concept cat-ELCCA Experiments. A) Titration of commercial Dicer. B) cat-ELCCA proof-of-concept with Dicer and inactivated Dicer. C) Evaluation of Dichlorofluorescein (DCF), guanosine diphosphate (GDP), and streptomycin in cat-ELCCA with either chemiluminescent (left) or Fluorescent (right) detection

3.2 cat-ELCCA Optimized for HTS

To adapt Dicer cat-ELCCA for high-throughput screening, several optimizations were made to reduce cost and improve upon the existing platform. It is well known that copper-catalyzed azide alkyne click chemistry (CuAAC) produces high yielding reactions with respect to small molecule synthesis, but what is less discussed is the low yields

associated with large biomolecular conjugations.[8] These low yields were observed with the pre-miRNA-HRP conjugate, prompting the exploration of other chemistries. In comparison to CuAAC, the Inverse Electron Demand Diels-Alder (IEDDA) reaction is significantly faster and does not require a catalyst.[13] While less common, IEDDA reactions have been reported for biomolecular conjugations and generally react a cyclopropene or transcyclooctene with a tetrazine.[14] However, there were no reports comparing IEDDA and CuAAC conjugations for RNA-Protein conjugates.

To determine if the improved rate constant and catalyst removal in the IEDDA reaction could lead to an increase in biomolecular conjugations, cyclopropene-NHS was synthesized and tested first due to its similar size to the alkyne modification.[15] Unfortunately, the cyclopropene had significant nonspecific labeling and was unstable, making it an ill-suited tool for bionconjugation reactions. As an alternative to the cyclopropene, the larger and faster transcyclooctene (TCO) was used. The pre-miRNA-TCO and -mTet conjugates were synthesized and tested with a rhodamine derivative in the same manner as the alkyne version. Similar yields were obtained when comparing the IEDDA and CuAAC reactions between RNA and rhodamine (Figure 3.5A and B, lanes 1, 4, and 7). Interestingly, the IEDDA reaction produced significantly higher yields compared to CuAAC when HRP was used instead of rhodamine, suggesting that the IEDDA reaction is better for large biomolecular conjugations (Figure 3.5B, lanes 3, 6, and 9). The 6-fold boost in yield translated to cat-ELCCA as a TCO-labeled RNA and tetrazine-labeled HRP produced much higher signal intensity than the respective alkyne and azide partners (Figure 3.5C and D). The increase of signal allowed for the reduction in the amount of RNA needed, significantly lowering the cost for HTS. Another notable

optimization was switching from a commercial source of Dicer to Dicer that was produced recombinantly in insect cells.[16] Importantly, the IEDDA modifications did not interfere with recombinant Dicer processing (Figure 3.5E). The combination of these optimizations led to a nearly ten-fold reduction in cost and increased the Z' to 0.69 (Figure 3.5F). Both of these optimizations were critical for the adaptation to HTS.
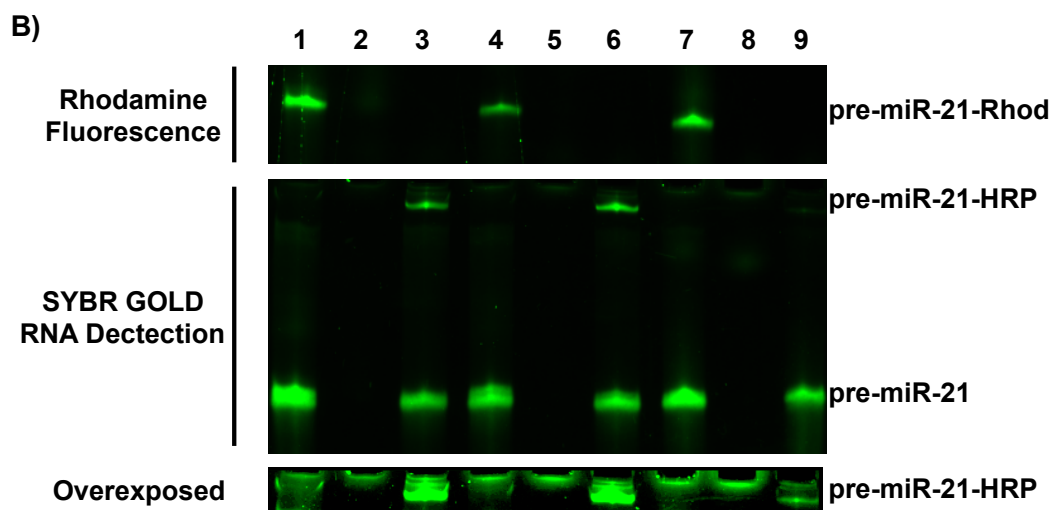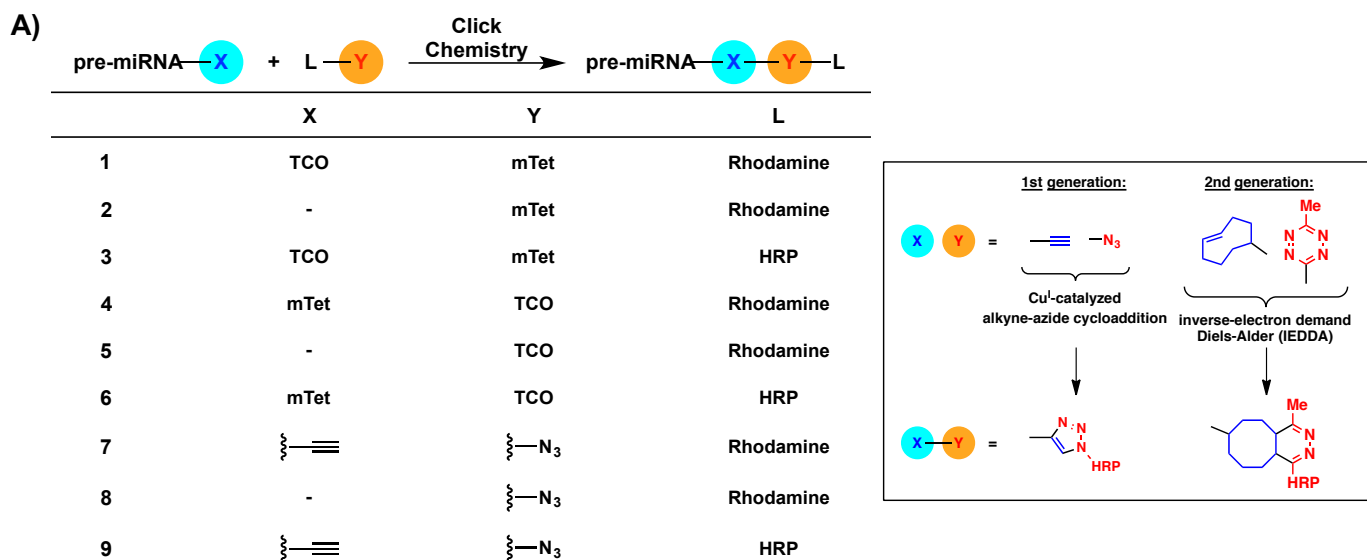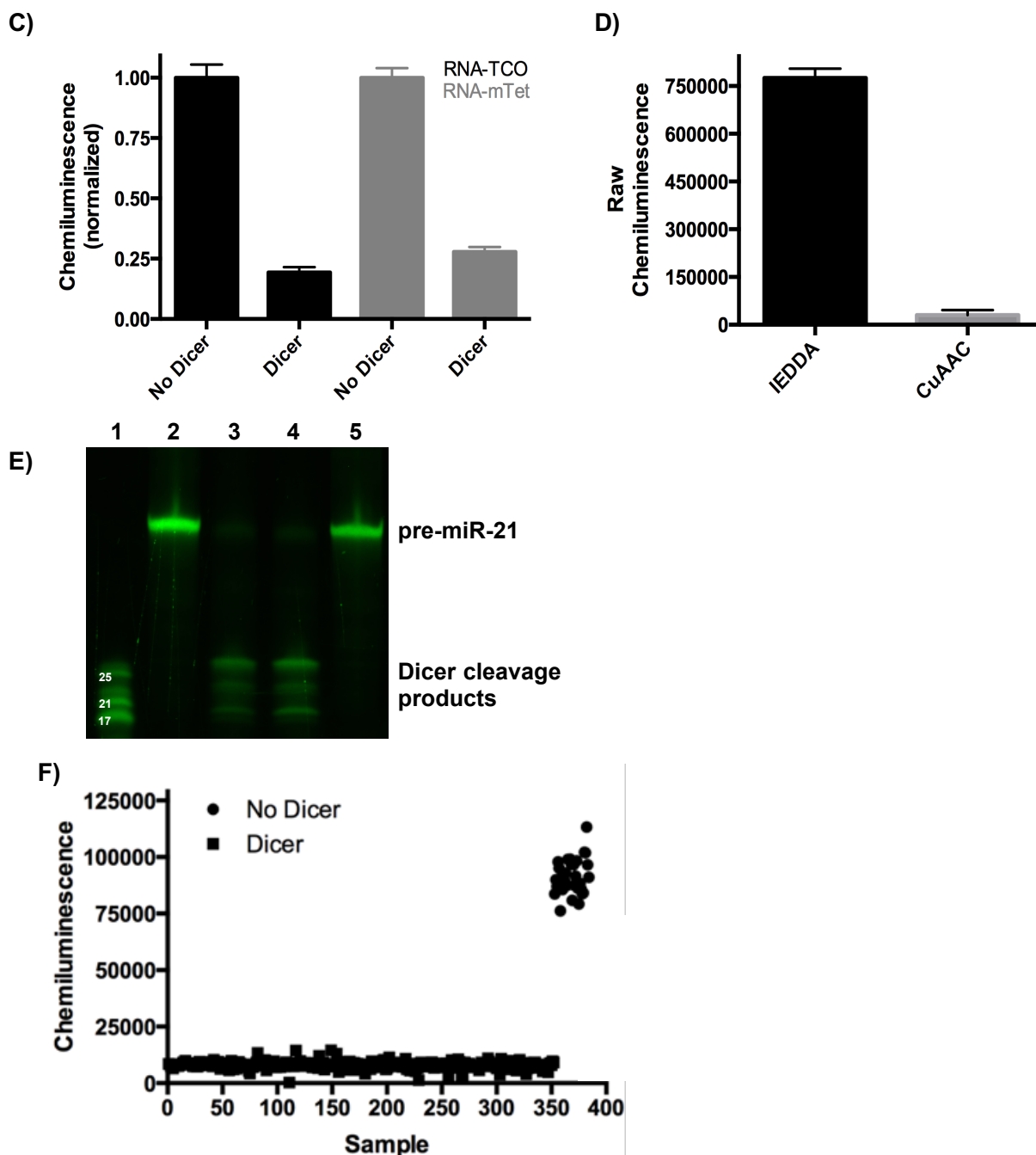
Figure 3.5. IEDDA cat-ELCCA Proof-of-Concept. A) General scheme and table of click reactions. B) pre-miRNAclick reactions between rhodamine (top) and HRP (middle). Overexposure of SYBR gold staining revealed the RNA-HRP conjugate for copper catalyzed click chemistry (bottom). Lanes are labeled according to A. C) Comparison between pre-miR-21 labeled with TCO and mTet in cat-ELCCA using the respectively conjugated HRP.D) Comparison between the IEDDA and CuAAC signal intensities in cat-ELCCA E) In Vitro digested pre-miR-21 substrates with recombinant Dicer. Lane 1 = miRNA ladder, lane 2 = pre-miR-21-TCO, lane 3 = pre-miR-21-TCO + Dicer, lane 4 = pre-miR-21-mTet + Dicer, lane 5 = pre-miR-21-mTet. F) 384- well plate HTS control plate with 352 wells treated with Dicer and 32 wells without.

3.3 Small Molecules Inhibitors of Dicer Dependent miRNA Maturation

After the optimizations to cat-ELCCA, a HTS campaign was conducted to identify inhibitors of pre-miR-21 maturation. In total, 48,127 small molecules and known drugs housed at the University of Michigan Center for Chemical Genomics (UMCCG) were tested at 25 μM for their ability to inhibit Dicer (Figure 3.6A). From this screen, 3.1% of molecules met our 5% inhibition criteria, affording 1,480 hits that were subsequently retested in triplicate to confirm activity. In addition to removing molecules that failed to repeat, compounds that were categorized as pan-assay interference compounds (PAINS) or generally reactive were also removed, resulting in 170 confirmed hit molecules.[17] To assess the potency of these molecules and evaluate their selectivity to our target, miR-21, over a control miRNA, let-7d, the hits were then tested in 8-point dose-response curves from 3.3–120 μM. Let-7d was used as a control counter screen because it too is a pre-miRNA hairpin loop, but consists of a completely different sequence (Figure 3.6B). Therefore, compounds that inhibit both pre-miR-21 and pre-let7-d maturation are likely either non-specific RNA binders or general Dicer inhibitors. Results from the comparison revealed that no compounds had high levels of selectivity for our target over the control; however, several compounds showed double digit micromolar IC_{50} values (Figure 3.6C).

A)

B)

pre-miR-21:   pre-let-7d:

```
      u        a       a       a     u    a              a                  c    uuagggcagggau
   agcuuauc  gacug  uguug  cugu    g       a          gagguaguagguug  auaguu                 u
                                        u                                                      u
   ucggguag  cugac  acaac  ggua    c       c          uuccgucguccagc  uaucaa                  u
  cug        -      c      -       -   u              ucu             a    uggaggaacaccc     g
```
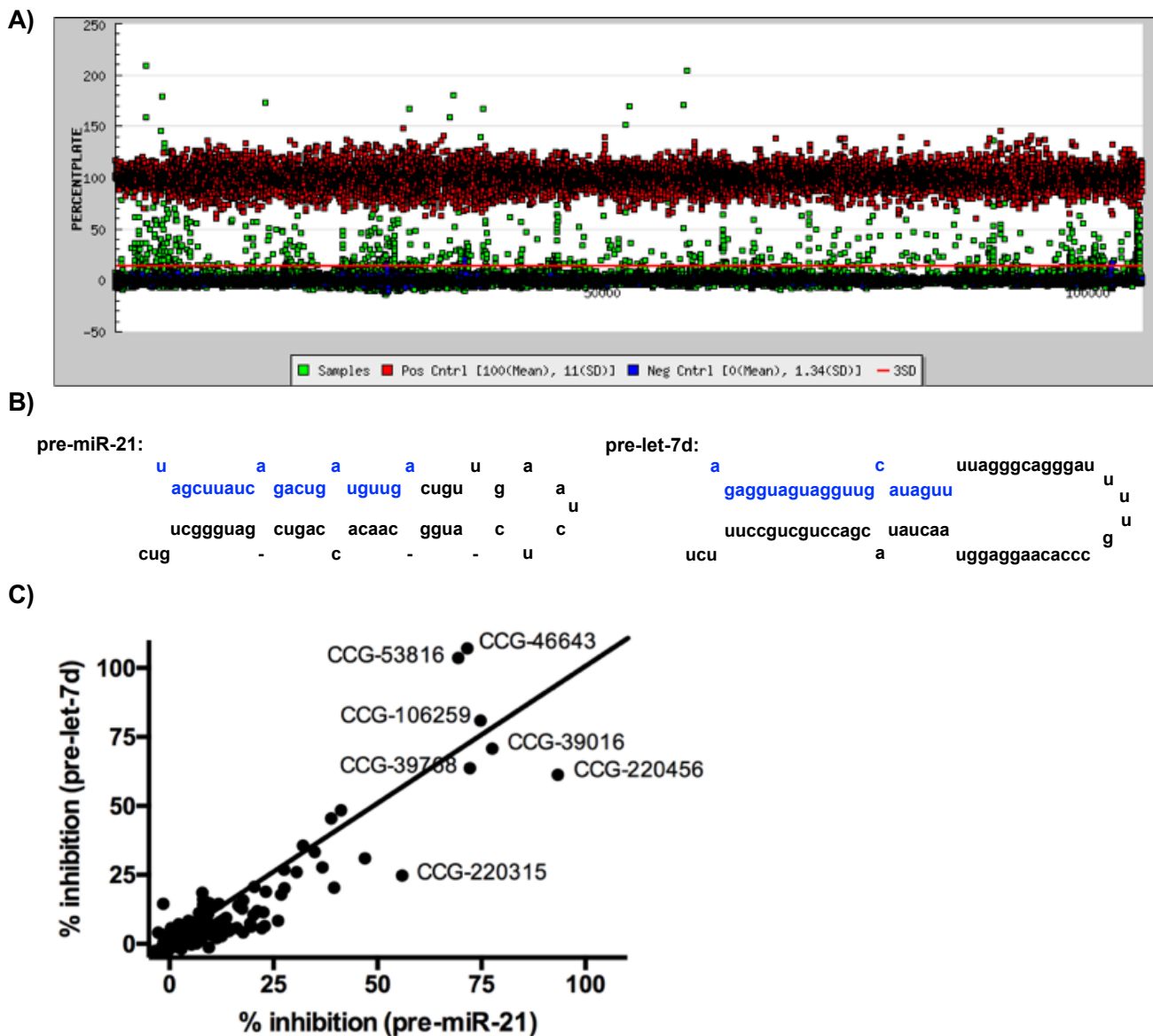
C)

Figure 3.6. Dicer Small Molecule HTS Data. A) High throughput screening campaign view with No Dicer, Dicer, and compound treated wells colored red, blue, and green respectively. B) Comparison between pre-miR-21 and pre-let-7d sequence and predicted secondary structure C) Two dimensional analysis of initial hits at 25 μM.

Currently, there is little information as to how to inhibit Dicer-dependent pre-miRNA maturation with small molecules. Additionally, many reported compounds have failed to be reproduced in other labs, including streptomycin in our hands (Figure 3.4C). Therefore, any insight that can be gathered to understand how molecules can be designed to reliably
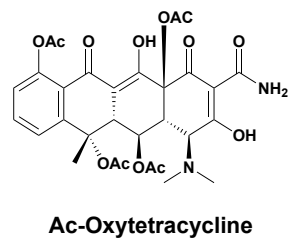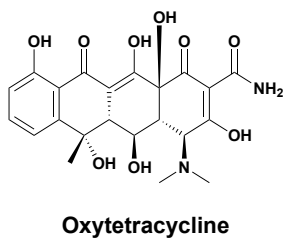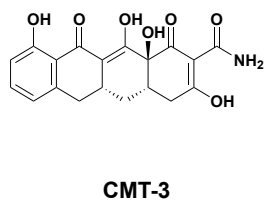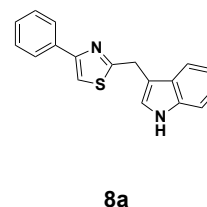
inhibit Dicer would be invaluable to the field. To this end, 8 of the hit compounds displaying varying degrees of selectivity and activity were synthesized in collaboration with another student in the lab, Jorge Sandoval, or purchased as solids, with seven showing some level of inhibition when freshly dissolved (Figure 3.7A and B).

Two of the eight compounds ordered, Methacycline and Meclocycline, are members of the tetracycline antibiotic family. What makes tetracyclines interesting for Dicer inhibition is their ability to bind both RNA and magnesium ions.[18] To explore the mechanism of inhibition for tetracyclines, several additional family members that were either located at the CCG or purchased were tested. Included in these tetracyclines was CMT-3, a tetracycline derivative that does not exhibit antibacterial activity presumably due to its inability to bind RNA.[19] The lack of RNA binding was confirmed in collaboration with Erin Gallagher by surface plasmon resonance (SPR). All of the tetracyclines, including CMT-3, show varying degrees of activity (Figure 3.7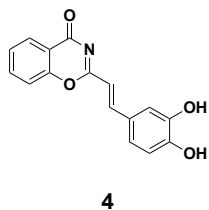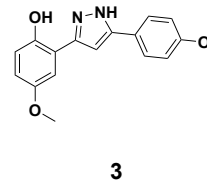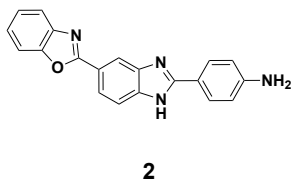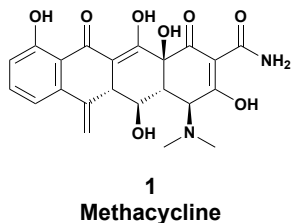C and D). The activity of CMT-3 suggests the metal-binding role is vital to tetracycline's activity, not their ability to bind RNA. To probe the metal-coordinating activity of tetracyclines, an acetylated oxytetracycline was designed and synthesized to prevent metal coordination in collaboration with Dr. Tanpreet Kaur. This acetylated oxytetracycline showed no activity (Figure 3.7E), further suggesting that metal coordination properties are essential for tetracycline-based inhibition of Dicer, not the RNA-binding properties. It is also important to note that the Dicer digest buffer contains over 5 equivalents of magnesium compared to the tetracycline, so the observed inhibition is unlikely due to metal sequestration.

Intriguingly, two of the other scaffolds contained a catechol moiety that is also known to bind magnesium.[20] The critical importance of the catechol was demonstrated

when a derivative lacking the catechol, 8a, was found to be inactive (Figure 3.7B). To rule out that catechols are nonspecific or interfere with our assay, two other catechol-containing molecules, Dopamine and 1,2-Dihydroxybenzene, were tested and both lacked the ability to inhibit Dicer. The above results suggest that molecules possessing magnesium-coordinating properties represent a new mechanism for Dicer inhibition that could be incorporated into new scaffolds to improve the reliability and reproducibility of Dicer inhibitors.

A)

1 Methacycline

2

3

4

5 Meclocycline

6

7

8

8a

CMT-3

Oxytetracycline

Ac-Oxytetracycline

B)

55
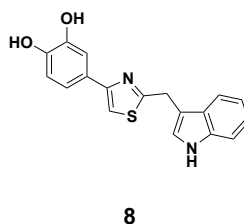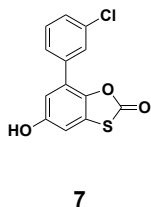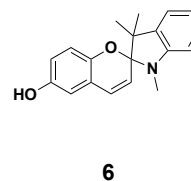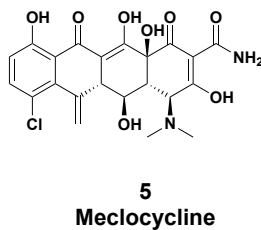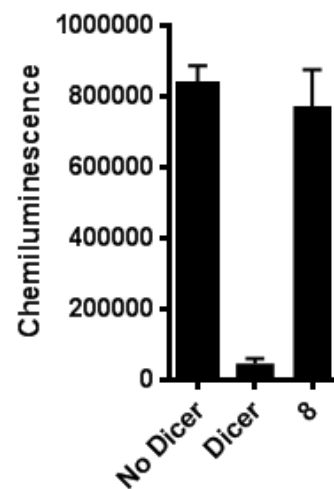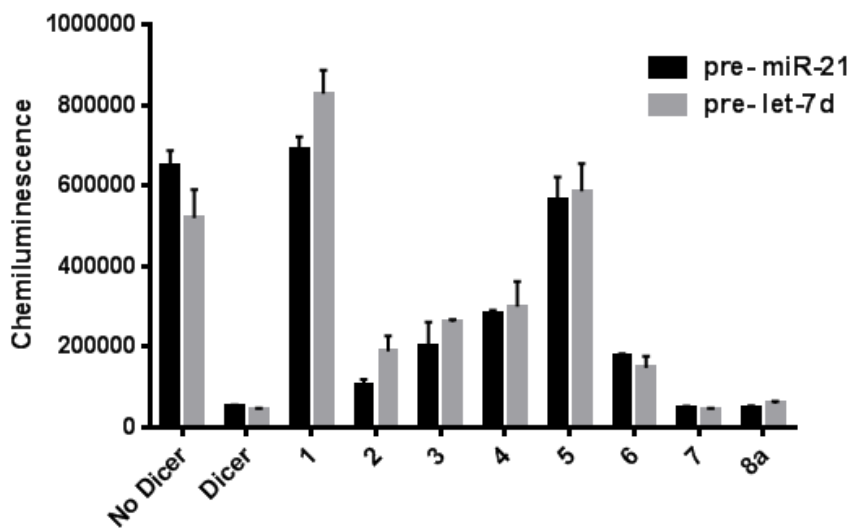
Figure 3.7. Small Molecule Dicer Hits. A) Compounds selected for further analysis. B) Selected compounds were tested at 1mM in cat-ELCCA. C) Various tetracycline derivatives tested at 1mM for Dicer inhibition in cat-ELCCA D) Measuring CMT-3 activity in cat-ELCCA at 1mM E) In Vitro digest of pre-miR-21-TCO comparing oxytetracycline and acetylated oxytetracycline at 1mM.

3.4 Identification of New RNA Binding Scaffolds

The results from the small molecule screen demonstrated that cat-ELCCA can be used to identify molecules capable of inhibiting Dicer-dependent maturation of pre-miRNAs. However, as with all screening libraries, the compounds at UMCCG are limited by known chemical space. Traditionally, natural products-based scaffolds have served as some of the most influential drugs used in the clinic. In an effort to identify new scaffolds that possessed RNA-binding properties that are either selective, or could have selectivity engineered into them, 32,301 natural product extracts (NPEs) were screened. Collected by Prof. David Sherman and colleagues, these e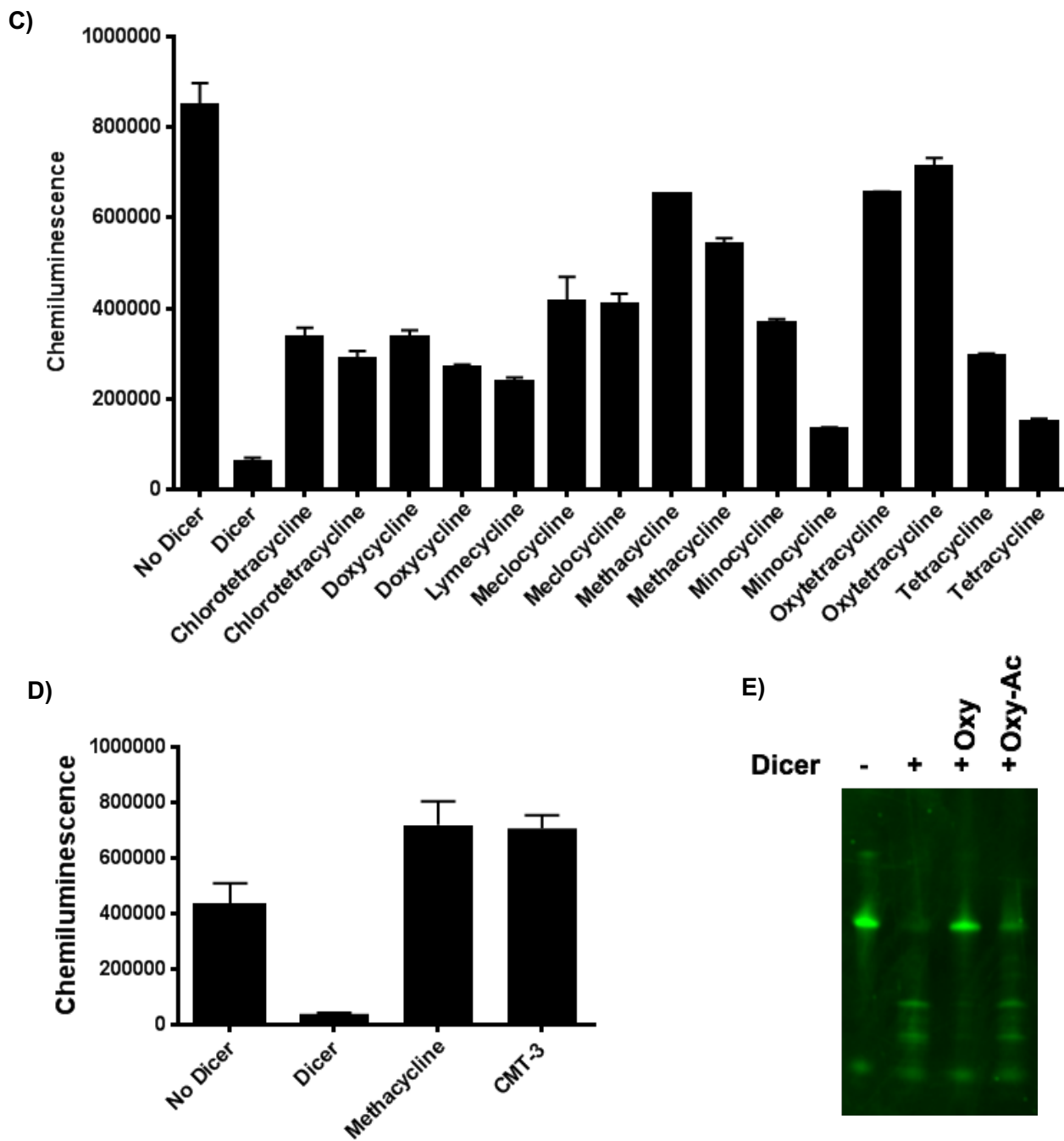xtracts originate from marine sediment samples containing actinobacteria and are housed at the UMCCG as crude extracts.[21] These extracts were tested in cat-ELCCA at 75 μg/mL for their ability to inhibit Dicer by at least 10%, affording 339 initial hits or a 1% hit rate. Hits were then validated in triplicate against both miR-21 and let-7 with some hits displaying varying degrees of selectivity and potency against both targets (Figure 3.8). While only noticeable in this experiment, it should be noted that a general trend toward the inhibition of miR-21 was observed. In addition to indicating binding selectivity for miR-21, this trend could also be explained by the different processing kinetics of Dicer substrates, as let-7 has properties that are consistent with faster processing then miR-21.[22] Instead of obtaining dose-response curves on impure extracts, 22 selected extracts were regrown in collaboration with the Sherman lab to determine if activity could be confirmed. These fresh extracts were again tested as single samples against both targets. After narrowing down the 22 extracts, two strains were selected in hopes of identifying a new RNA-binding scaffold.

Since natural product extracts contain numerous compounds, they require several rounds of purifications and testing. These experiments are currently underway in collaboration with Dr. Ashu Tripathi and Jorge Sandoval. Preliminary results suggest that one of the the hit scaffolds is a peptide derivative. A series of NMR and high resolution mass spectrometry experiments are currently underway to assign the final structure. Additionally, in collaboration with Erin Gallagher, the hit scaffold was tested by SPR and found to bind pre-miR-21 with nanomolar affinity. Taken together this data shows that natural products contain unique chemical space to inhibit pre-miRNA maturation.

A)

B)
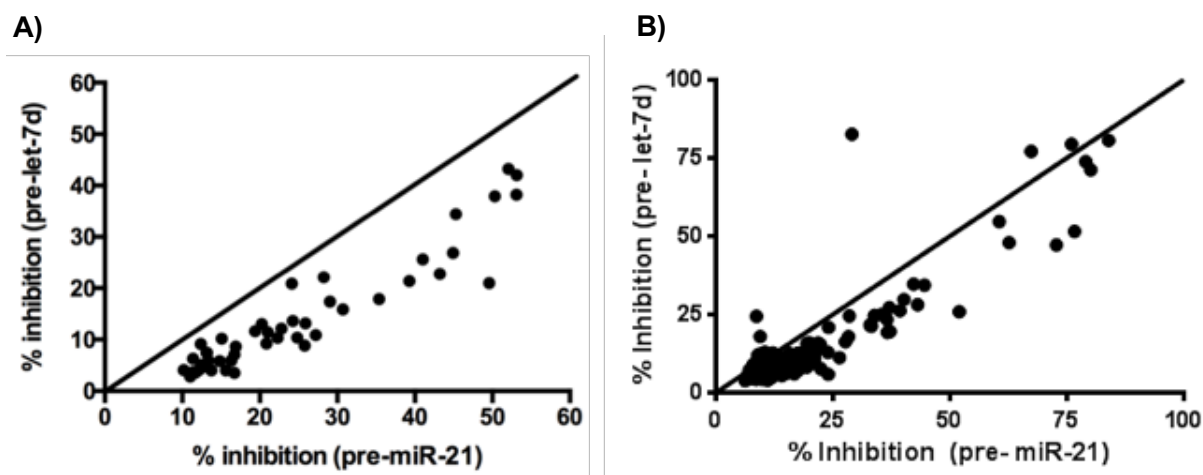
Figure 3.8. Dicer Natural Product HTS Data. A) Two dimensional analysis of the average initial hits tested in triplicate B) Two dimensional analysis of selected extracts tested as single samples.

3.5 Conclusion

The development of cat-ELCCA for Dicer maturation has been instrumental in discovering new molecules capable of regulating miRNA biology. Even though no lead compounds selective for miR-21 were discovered, the insights gained from the magnesium-coordinating molecules provide a novel mechanism to incorporate into RNA-binding molecules to inhibit Dicer-based processing. Additionally, through the successful demonstration of natural product screening, cat-ELCCA represents a new way to identify

new scaffolds for the inhibition of Dicer dependent miRNA maturation. This is particularly important due to the high prevalence of chromophoric and fluorogenic compounds that interfere with most assays.

3.6 Methods

pre-miR-21 RNA Sequence:

The following sequence was ordered from Dharmacon:

5'-Biotin-(18-atom spacer; hexaethylene glycol)-UAGCUUAUCAGACUGAUGUUGACUGUUGAA-(5-aminoallyl uridine)-CUCAUGGCAACACCAGUCGAUGGGCUGUC-3'

pre-let-7d RNA Sequence:

The following sequence was ordered from Dharmacon:

5'-Biotin-(18-atom spacer; hexaethyleneglycol)-AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAGGGA-(5-aminoallyl uridine)-UUUGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUCU-3'

Preparation of RNA-Click Conjugate:

The NHS-ester of 4-pentynoic acid, TCO, and mTet were purchased or synthesized following previously reported methods and was used without further purification. To generate the cat-ELCCA substrate, RNA (5.0 μL of 1.0 mM stock in 100 mM, pH 8 phosphate buffer; 5.0 nmol final) was incubated with NHS ester (5.0 μL of 10 mM stock in DMSO; 50 nmol) for 1 h at 25 °C. The RNA conjugate was purified by precipitation with

sodium acetate (1.1 μL of 3M solution at pH 5.2) and cold ethanol (40 μL), followed by centrifugation at 14,000 RPM for 40 min at 4°C. RNA was stored long-term as a 1.0 mM stock (100 mM phosphate buffer, pH 8) at -80 °C.

Preparation of HRP-N$_3$:

HRP-N$_3$ was prepared following an established procedure and stored at 4 °C in phosphate buffer (100 mM, pH 7.0).[23] Q-TOF HPLC-MS confirmed the coupling of 4 azides per molecule of HRP. HRP mass: 43261.6294, HRP-N$_3$ mass: 43365.1259

Preparation of HRP-TCO and HRP-mTet:

2.5 mg HRP was dissolved in 185.8 μL PBS (100 mM phosphate buffer, pH 7.0, 150 mM NaCl) and mixed with 14.2 μL 100 mM TCO-PEG4-NHS or mTet-NHS dissolved in DMSO. The mixture was gently shaken at room temperature for 3 h then exchanged using a microcentrifuge concentrator into PBS to remove unreacted NHS esters and DMSO. HRP-TCO and HRP-mTet were stored at 4 °C.

Rhodamine-amine (RHOD-NH$_2$):
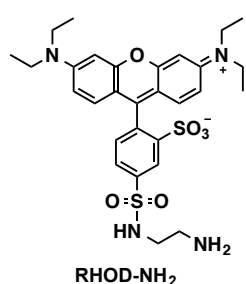
Lissamine rhodamine (0.25 mmol) was dissolved in 10 mL anhydrous DMF under N$_2$ followed by the addition of *N*-Boc-ethylenediamine (0.375 mmol) and triethylamine (1.25 mmol). The reaction was stirred at 25 °C overnight. The mixture was extracted with ethyl acetate, washed with saturated sodium bicarbonate, and the organic layer was dried *in vacuo* overnight. The resulting crude residue was purified by HPLC. The Boc group was

removed by addition of 80% trifluoroacetic acid in dichloromethane for 1 h at 25 °C. The final product was concentrated *in vacuo* and dissolved in DMSO. RHOD-NH$_2$ *m/z* calc. [M+H]$^+$ 601.2149, found 601.2145.

Rhodamine-TCO (RHOD-TCO) and Rhodamine-mTet (RHOD-mTet):

RHOD-NH$_2$ (1.66 μmol) was mixed with either TCO-PEG4-NHS or mTet-NHS (1.66 μmol) in DMSO followed by the addition of triethylamine (2 μL). The reaction was allowed to proceed overnight at 25 °C. Products were confirmed by analytical HPLC and mass spectroscopy and used as is. RHOD-TCO *m/z* calc. [M+H]$^+$ 1000.4406, found 1000.4384; RHOD-mTet *m/z* calc. [M+H]$^+$ 813.2847, found 813.2844.

RNA-Rhodamine Click Chemistry:

RNA (500 nM final) was reacted with Azide-fluor 488, or RHOD-TCO, or RHOD-mTet (1 μM final) using standard click chemistry conditions 100mM phosphate buffer pH7 (Supplemented with 100 μM CuSO$_4$, 500 μM THPTA, 5.0 mM sodium ascorbate for CuAAC). After incubation for 2 h at 25 °C, the RNA was diluted 2x with RNA loading dye (95% Formamide, 0.02% SDS, 0.02% Bromophenol blue, 0.01% Xylene Cyanol, 1mM EDTA) and analyzed by 10% TBE Urea gel eletrophesis.

Quantification of Immobilized RNA:

The wells of a streptavidin-coated 384-well plate were washed 2x with 50 μL of 100 mM phosphate buffer (pH 7.0). Immobilization was then carried out by adding 10 μL of 500 nM RNA in 100 mM phosphate buffer (pH 7) to the wells, which were agitated overnight

at 4 °C. Immobilization efficiency was determined by analyzing the incubated solution by gel-electrophoresis followed by quantification by densitometry.

Dicer Purification:

Dicer was prepared as reported; however, the enzyme was instead dialyzed overnight and stored at -20 °C in 20 mM Tris pH 7.5, 100 mM NaCl, 1.0 mM $MgCl_2$, 50% glycerol, and 0.1% Triton X-100.[16]

Dicer Digestion:

Solution digests were carried out in 10-μL volume. **RNA-X** (500 nM final) was treated with Dicer (1.0 µL, 1.3mg/ml) in buffer (20 mM Tris-HCl, pH 7.4, 12 mM NaCl, 2.5 mM $MgCl_2$, 40 U/mL RNase Out, 1.0 mM fresh DTT) at 37 °C for 3 h. Digests were analyzed using a 12.5% TBE-Urea gel and visualized using SYBR Gold.

HTS Assay Protocol

Black, standard capacity streptavidin-coated 384-well plates (Pierce 15407) were first washed with 50 μL of sodium phosphate buffer (100 mM, pH 7.0; PB7) three times using a Biotek 405 ELX plate washer. Subsequently, 5 μL of biotinylated pre-miRNA substrate (500 nM final) was dispensed into the plate using a Multidrop Combi Reagent Dispenser (Thermo Scientific). Plates were then centrifuged for 1 min at 1,000 RPM ($223 \cdot g$), sealed with plate tape, and incubated overnight at 4 °C. The following morning, plates were washed three times with 50 μL of PB7, followed by the addition of 5 μL of Dicer digest buffer (20 mM Tris, 12 mM NaCl, 2.5 mM $MgCl_2$, 1 mM fresh DTT, and 4.5% DMSO) and

centrifugation. Compounds (50 nL of 5 mM DMSO stock, 25 µM final) were then added into the sample wells using a Sciclone (Caliper) liquid handler with V&P pintool; the same volume of DMSO was added to the control wells. The plates were incubated at 25 °C for 15 min before addition of 5 µL of digest buffer containing 217 µg/nL Dicer (108 µg/mL Dicer, 5% glycerol and 0.01% Triton X-100 final). For the positive control wells, digest buffer without Dicer was added. The plates were centrifuged again and resealed before being placed in a 37 °C incubator for 5 h. After Dicer cleavage, plates were washed three times with 50 µL of PB7. mTet-HRP in PB7 (10 µL, 750 nM final) was then dispensed into each well. The plates were subsequently centrifuged, sealed, and incubated at 25 °C for 2 h. Plates were then washed three times with 50 µL of wash buffer (2 mM imidazole, 260 mM NaCl, 0.5 mM EDTA, 0.1% Tween-20, pH 7.0), followed by washing three additional times with 50 µL of PB7. Finally, SuperSignal West Pico (25 µL; Pierce) was added, the plates were incubated at 25 °C for 5 min, and chemiluminescence signal was detected using a PHERAstar plate reader using LUM plus module (BMG Labtech).

Cat-ELCCA Protocol (By Hand):

Same as "HTS Assay Protocol" with the following modifications: All washing and dispensing was done by hand with a pipette, plates were not centrifuged following additions, and chemiluminescence was detected on a Biotek Cyation3.

Compound Libraries

Compounds screened were housed at the University of Michigan Center for Chemical Genomics (CCG). For the primary screen, 47,130 compounds from the following

collections were used: Sigma LOPAC library of pharmacologically active compounds (1,280), Prestwick library of approved drugs (1,280), ChemDiv 100K library (21,120), Maybridge MB24K library (23,552), and UM Chemistry library (895). Additionally, a library of 32,301 natural product extracts (NPE) library was also tested.[9] Compounds were tested at 25 μM in the primary and confirmation screens using 5 mM DMSO stocks. Concentration response curves were generated over 8 points (1.67-fold serial dilution) from 3.3–120 μM using 5 mM DMSO stocks; however, compounds were first dispensed with a Mosquito x1 (TTP Labtech) into polypropylene 384-wells plates (Greiner 784201), and subsequently diluted with Dicer digest buffer (15 μL) before addition of diluted compound into the pre-miRNA-immobilized plate (5 μL). NPEs were tested at 75 μg/mL in the primary and confirmation screens using 15 mg/mL stocks.

Data Analyses

HTS data was monitored and analyzed using MScreen.[11] Small molecules were considered as initial hits if they exhibited ≥5% inhibition by plate based on the negative controls. For the NPEs, this threshold was raised to ≥10% inhibition by plate based on the negative controls. Potential hits meeting these criteria (1,480 small molecules and 339 NPEs) were confirmed by rescreening in triplicate. Compounds showing inhibition at ≥3SD by plate from the negative controls were considered as confirmed hits and analyzed in concentration response curves in duplicate (170), excluding the NPEs, which underwent more stringent analysis to select those for regrowth. Average percent inhibitions by plate at 120 μM (small molecules) and 75 μg/mL (NPEs) were determined from sample and positive control values normalized to the negative control. All data was

analyzed using GraphPad Prism version 6.0c for Mac OS X (GraphPad Software,

www.graphpad.com).

3.7 Copyright

The work in this chapter was reproduced in part from 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 chemistry* **26,** 19–23 (2015), Lorenz, D. A. & Garner, A. L. A click chemistry-based microRNA maturation assay optimized for high-throughput screening. *Chem. Commun.* **52,** 8267–8270 (2016), 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 Discov* 2472555217717944 (2017). doi:10.1177/2472555217717944, and Lorenz, D. A. & Garner, A. L. in *RNA Therapeutics* (ed. Garner, A. L.) 79–110 (Springer International Publishing, 2018).

3.8 References

1.   Garner, A. L. & Janda, K. D. cat-ELCCA: a robust method to monitor the fatty acid acyltransferase activity of ghrelin O-acyltransferase (GOAT). *Angewandte Chemie (International ed. in English)* **49,** 9630–9634 (2010).

2.   Garner, A. L. & Janda, K. D. A small molecule antagonist of ghrelin O-acyltransferase (GOAT). *Chem. Commun.* **47,** 7512–7514 (2011).

3.   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 chemistry* **26,** 19–23 (2015).

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

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 Discov* 2472555217717944 (2017). doi:10.1177/2472555217717944

6.   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.* **19,** 763–769 (2017).

7.   Lequin, R. M. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). *Clin. Chem.* **51,** 2415–2418 (2005).

8. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie (International ed. in English)* **40,** 2004–2021 (2001).

9. Tian, Y., Simanshu, D. K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V. N. & Patel, D. J. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular Cell* **53,** 606–616 (2014).

10. Krichevsky, A. M. & Gabriely, G. miR-21: a small multi-faceted RNA. *J. Cell. Mol. Med.* **13,** 39–53 (2009).

11. Torimura, M., Kurata, S., Yamada, K., Yokomaku, T., Kamagata, Y., Kanagawa, T. & Kurane, R. Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Anal Sci* **17,** 155–160 (2001).

12. Zhang, J., Chung, T. & Oldenburg, K. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4,** 67–73 (1999).

13. Lang, K. & Chin, J. W. Bioorthogonal Reactions for Labeling Proteins. *ACS chemical biology* **9,** 16–20 (2014).

14. Knall, A.-C. & Slugovc, C. Inverse electron demand Diels–Alder (iEDDA)-initiated conjugation: a (high) potential click chemistry scheme. *Chem. Soc. Rev.* **42,** 5131–5142 (2013).

15. Patterson, D. M., Nazarova, L. A., Xie, B., Kamber, D. N. & Prescher, J. A. Functionalized Cyclopropenes As Bioorthogonal Chemical Reporters. *Journal of the American Chemical Society* **134,** 18638–18643 (2012).

16. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 512–517 (2008).

17. Baell, J. B. & Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **53,** 2719–2740 (2010).

18. Guerra, W., Silva-Caldeira, P. P., Terenzi, H. & Pereira-Maia, E. C. Impact of metal coordination on the antibiotic and non-antibiotic activities of tetracycline-based drugs. *Coordination Chemistry Reviews* **327-328,** 188–199 (2016).

19. Golub, L. M., McNamara, T. F., D'Angelo, G., Greenwald, R. A. & Ramamurthy, N. S. A non-antibacterial chemically-modified tetracycline inhibits mammalian collagenase activity. *J. Dent. Res.* **66,** 1310–1314 (1987).

20. Kim, Y. J., Wu, W., Chun, S.-E., Whitacre, J. F. & Bettinger, C. J. Catechol-mediated reversible binding of multivalent cations in eumelanin half-cells. *Adv. Mater. Weinheim* **26,** 6572–6579 (2014).

21. Lowell, A. N., Santoro, N., Swaney, S. M., McQuade, T. J., Schultz, P. J., Larsen, M. J. & Sherman, D. H. Microscale Adaptation of In Vitro Transcription/Translation for High-Throughput Screening of Natural Product Extract Libraries. *Chem Biol Drug Des* **86,** 1331–1338 (2015).

22. Zhang, X. & Zeng, Y. The terminal loop region controls microRNA processing by Drosha and Dicer. *Nucleic Acids Research* **38,** 7689–7697 (2010).

23. van Dongen, S. F. M., Teeuwen, R. L. M., Nallani, M., van Berkel, S. S., Cornelissen, J. J. L. M., Nolte, R. J. M. & van Hest, J. C. M. Single-step azide introduction in proteins via an aqueous diazo transfer. *Bioconjugate chemistry* **20,** 20–23 (2009).

Chapter 4

Targeting Lin28 with Small Molecules

One mechanism that life has developed to ensure the proper regulation of RNAs is the use of RNA-binding proteins (RBPs) to regulate RNA function (Chapters 1 and 2). Over 1,500 RBPs are predicted in *Homo sapiens* and their targets are quickly being elucidated through several technologies such as cross-linking immunoprecipitation (CLIP).[1,2] As our knowledge of the biological implications of RBPs continues to increase, many RBPs have already been linked to or deemed clinically relevant targets.[3] Unfortunately, there are insufficient platforms to enable the discovery of small molecule probes and potential drug candidates for these clinically relevant RBPs. Based on the successful implementation and demonstration of Dicer cat-ELCCA in Chapter 3, cat-ELCCA would serve as an ideal platform to fill this technological void for facilitating the discovery of RNA-Protein interaction (RPI) modulators.[4,5] In addition to the standard benefits of cat-ELCCA previously demonstrated, the modular approach of cat-ELCCA lacks the requirement of structural information and the need to incorporate radioactivity, which are necessary for most current RPI assays.[6-8]

4.1 Lin28 cat-ELCCA Assay

To establish cat-ELCCA as a suitable assay for RPIs, the model system Lin28-Let-7 was chosen due to its biological significance and well-characterized interaction.[9] Similar to other RBP assays, a site-specific labeling technique was necessary for cat-ELCCA to avoid labeling residues that could be required for RNA binding. This site-selective labeling was designed to be achieved through the generation of a Halotag(HT)-RBP fusion protein as shown in Figure 4.1. Halotag was engineered from a haloalkane dehalogenase to use a bioorthogonal haloalkane suicide ligand to form a covalent bond with its active site.[10] Thus, a Halotag fusion protein is an ideal labeling technique for cat-ELCCA because it allows the platform to be readily adapted to any protein-of-interest without having to generate site-specific mutants. Additionally, there are numerous reports of labeling either protein termini with HT without affecting protein function, further demonstrating that this strategy would be well-suited for an assay platform like cat-ELCCA.[11] There are two key differences from the previous Dicer cat-ELCCA. First, the protein was chosen to be immobilized because of the poor immobilization efficiency of RNA that was demonstrated in Chapter 3. In fact, the assay did not produce satisfactory results when the RNA was immobilized and click handle incorporated into the HT (data not shown). Second, compounds that disrupt the RPI will result in reduced signal, in contrast to the Dicer assay, as reduced binding of RNA will lead to reduced HRP retention.
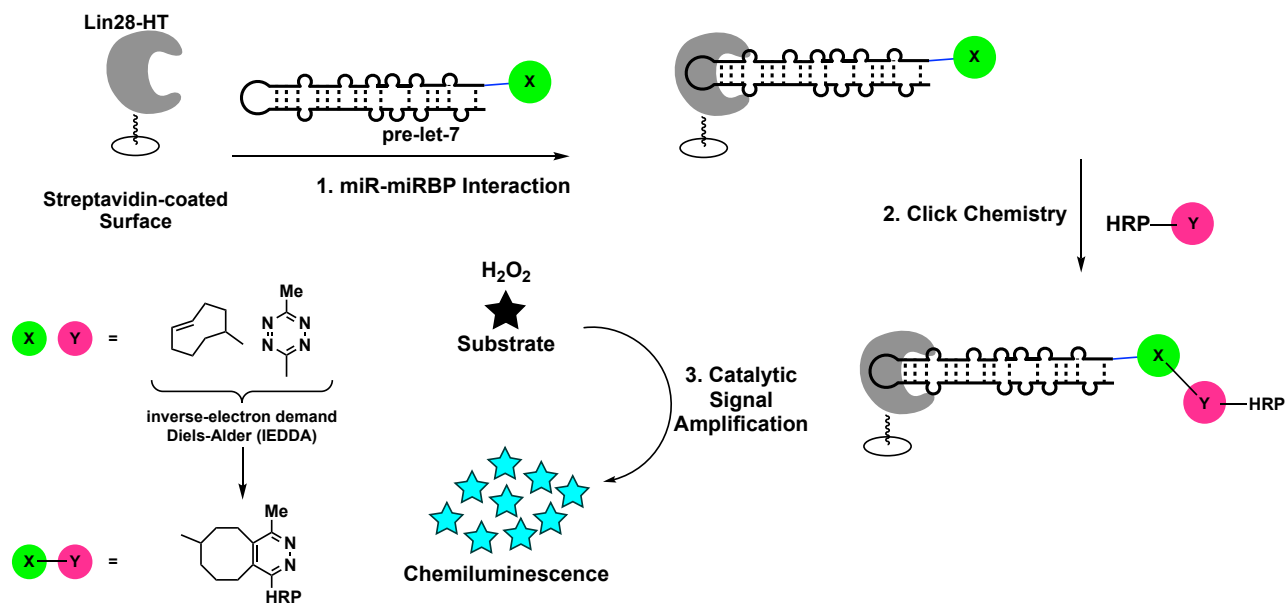
Figure 4.1. **RBP cat-ELCCA Scheme.**

Initial tests were designed to confirm the individual components could be generated and would function properly. First was the generation of the Lin28-HT fusion protein. The murine homolog of Lin28a, which has 97% homology to *Homo sapiens*, was fused to the C-terminus of HT and expressed recombinantly in *E. coli*. Following purification, the Lin28-HT protein was incubated with a biotin-chloralkane ligand to allow for the site-selective biotinylation (Figure 4.2A).[12] The presence of Lin28a, HT, and biotin were confirmed by Western blot analysis (Figure 4.2B). This fusion protein was then evaluated in an electrophoretic mobility shift assay (EMSA) to ensure that the fusion and biotin tag did not interfere with labeled Let-7 substrate binding. As seen in Figure 4.2C, the fusion protein yielded a dissocaitaion constant of approximately 200 nM, which is similar to other reported values.[9] The presence of the 5' trans-cyclooctene (TCO)-labeled RNA was confirmed by a reaction with a mTetrazine-Rhodamine and analyzed by gel electrophoresis using the same method from the Dicer assay in Chapter 3.
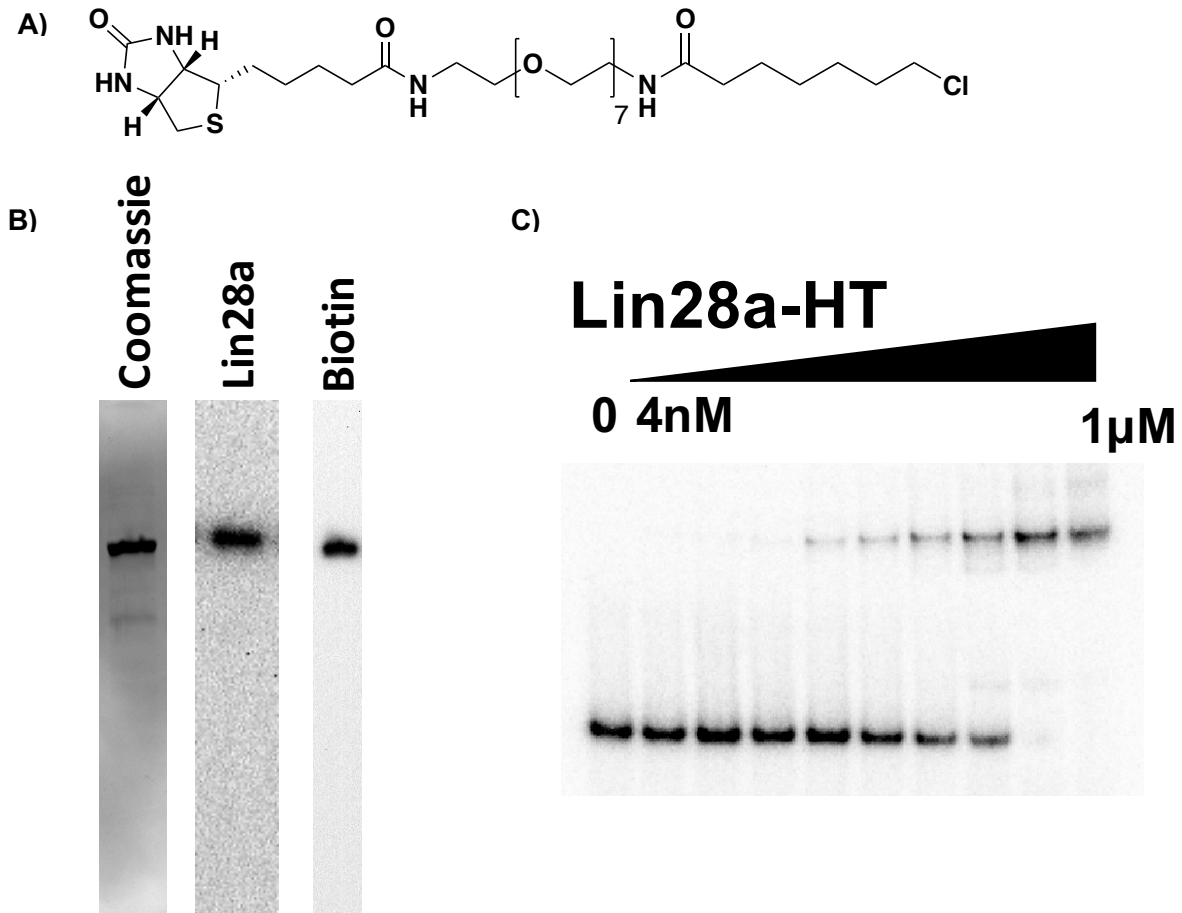
Figure 4.2. Generation of Lin28 Fusion Protein. A) Structure of biotin-peg7-Chloro ligand B) Coomassie stain and western blots for Lin28a and Biotin to confirm protein expression and labeling. C) Electrophoretic mobility shift assay of ^{32}P labeled pre-let7d-tco with Lin28a-HT-Biotin protein. Concentration tested from left to right: 0, 4, 8, 16, 31, 63,125, 250, 500, 1,000 nM.

With all of the components in hand, proof-of-concept experiments were preformed to test the RPI cat-ELCCA. The first of these experiments was to determine if the signal intensity would increase as more RNA was used. As can be seen in Figure 4.3A, the assay preformed as expected with an increase of signal proportional to the amount of pre-let-7 used. To further evaluate the assay, two different unlabeled RNAs, pre-let-7d and pre-miR-21 were used as competitors of the pre-let7 signal. Figure 4.3B shows that the signal decreased as unlabeled pre-let-7 was added; however, the same decrease

was not observed with pre-miR-21. This was an expected result because pre-miR-21

should not bind to Lin28a.[13] Finally, the assay was evaluated for its suitability for HTS by

determining the assay's reproducibility using automated equipment. Using 192 wells of

positive and 192 wells of negative controls, the assay yielded an excellent Z' value of 0.59

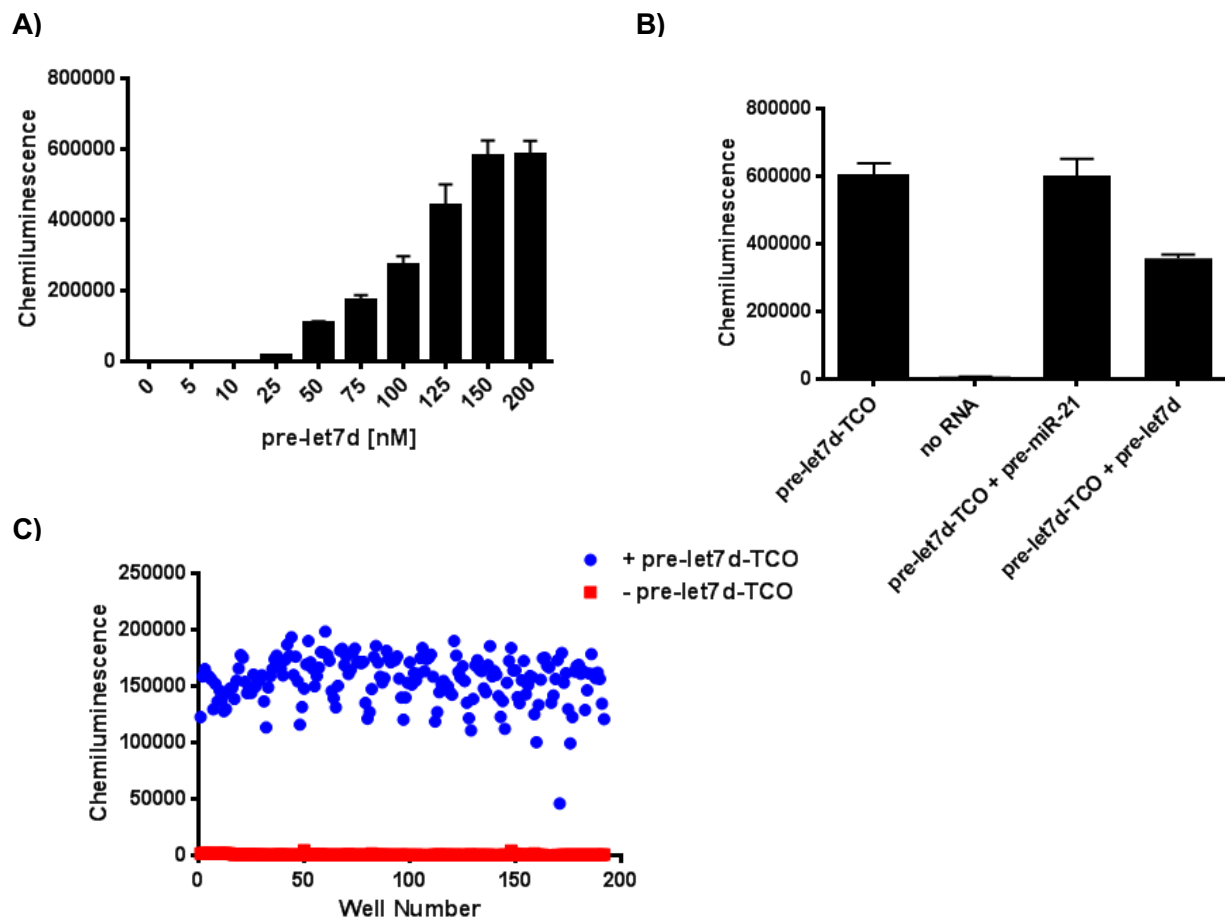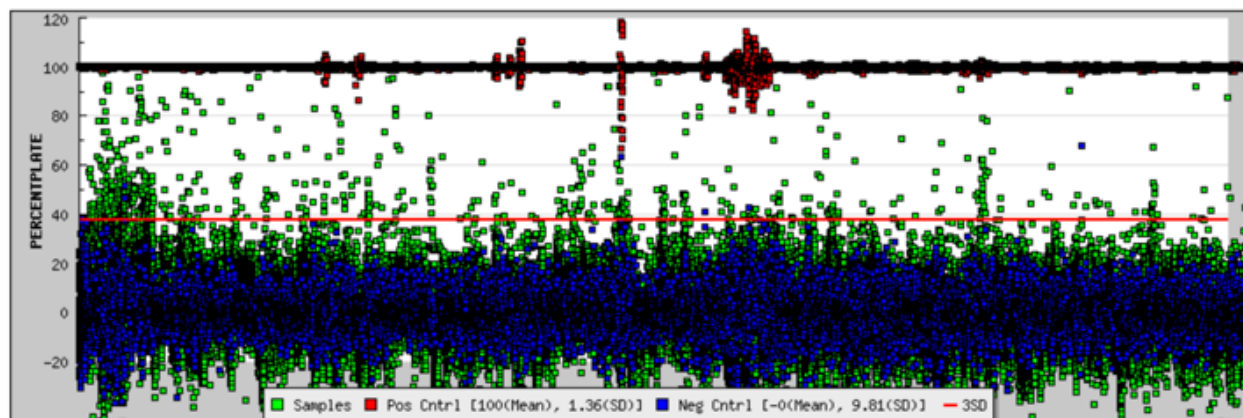and a signal-to-background ratio of 275 (Figure 4.3C).[14]

Figure 3. Lin28 cat-ELCCA Proof-of-Concept. A) Titrating labeled pre-let7d-TCO in RBP cat-ELCCA. B) Cat-ELCCA showing competition with 500 nM unlabeled pre-let7d and pre-miR-21. C) Preliminary Z' calculation data from 192 positive and 192 negative controls
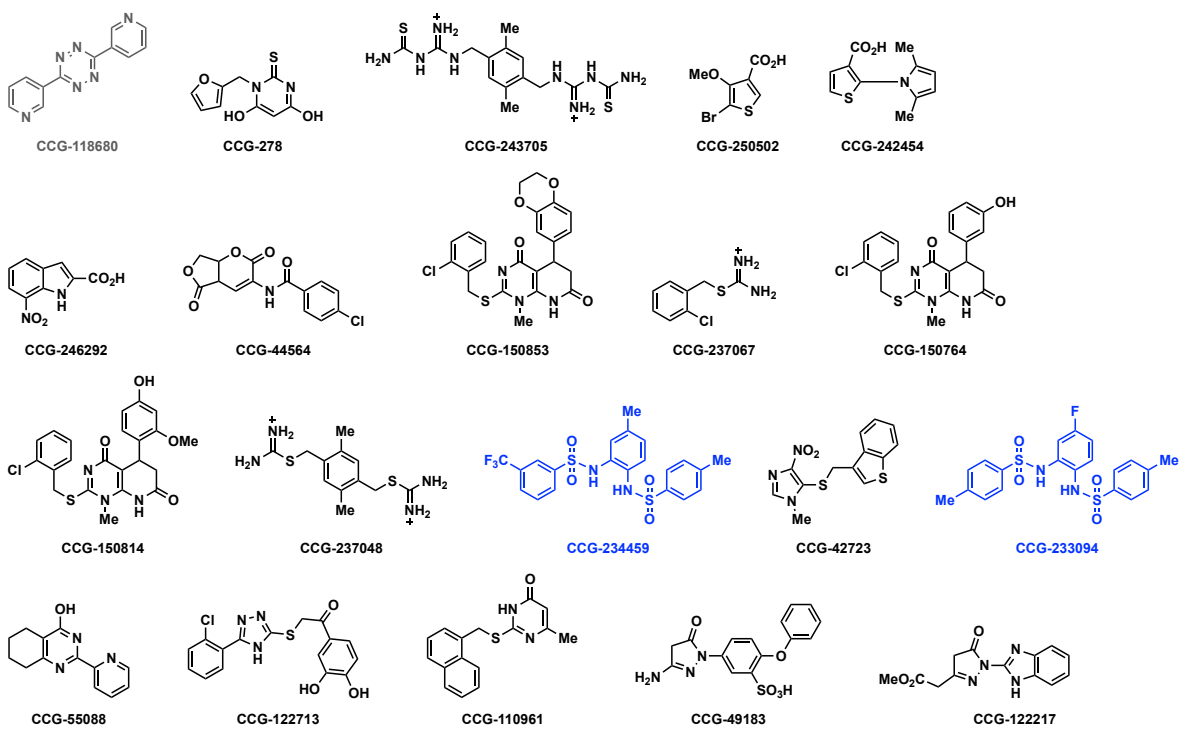
4.2 High-Throughput Screening

      With the assay in hand, 126,157 small molecules were screened at the University of Michigan Center for Chemical Genomics (UMCCG) in collaboration with Samuel Kerk. As can be seen from the campaign view in Figure 4.4A, the assay performed as expected with clear separation of controls and a good distribution of hits using 25 μM of compound. In total, 1,468 initial hits were identified using 25% inhibition as a cutoff. These hits were then replated and tested in triplicate with 181 compounds confirming at least 3 standard deviations from the control for a minimum of 2 out of 3 times. The large drop in repeating compounds that was observed can be partially attributed to errors in the automated liquid dispenser causing the occasional row of false positives. Hits that repeated were then tested in a dose response curve from 3.3–120 μM, with 136 compounds showing dose-dependent inhibition. These 136 compounds afforded a final hit rate of 0.11%. From these 136 compounds, 20 were purchased as fresh solids with 10 showing activity in cat-ELCCA (Figure 4.4B and C).

A)

B)

CCG-118680   CCG-278   CCG-243705   CCG-250502   CCG-242454

CCG-246292   CCG-44564   CCG-150853   CCG-237067   CCG-150764

CCG-150814   CCG-237048   CCG-234459   CCG-42723   CCG-233094

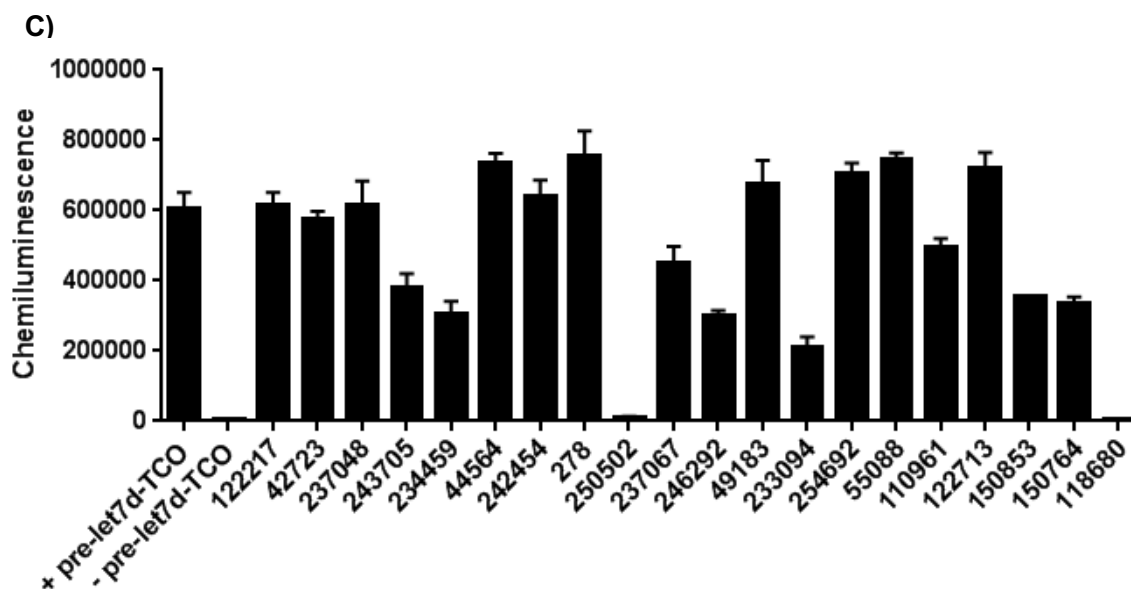CCG-55088   CCG-122713   CCG-110961   CCG-49183   CCG-122217

74

C)

Figure 4. Lin28 HTS and Preliminary Hits. A) Campaign view from lin28 HTS using cat-ELCCA. B) Structures of 20 compounds purchased with lead scaffold highlighted in blue. C) Fresh compounds tested in RPI cat-ELCCA at 100uM

4.3 Structure Activity Relationship

Of the ten active compounds, **1** and **2** shared a similar sulfonamide scaffold that exhibited selectivity when evaluated using the UMCCG and PubChem databases, only being classified as a hit in 1/11 and 1/9 assays, respectively (Figure 4.5A). To further validate this new scaffold, an EMSA was performed to confirm the inhibitory properties of **1** and **2** in a different assay format. Using both EMSA and cat-ELCCA titrations, it was determined that both **1** and **2** have low micromolar IC_{50} values (Figure 4.5B and C). The minor differences in values are attributed to the different concentrations of RNA and protein used in each assay, 500 and 200 nM, respectively. Encouraged by these results, a structure-activity relationship (SAR) campaign was started by ordering 20 additional sulfonamide derivatives (Figure 4.5A). While many compounds had similar IC_{50} values,

none were more potent than the original hit **1** (Figure 4.5A). However, these compounds still provided valuable SAR information. The requirement of ortho substitution around the center sulfonamide ring was highlighted by the inactivity of compounds **3** and **4**, as substitution at either the meta or para positions resulted in no activity (Figure 4.5B). Additionally, compounds **5** and **6** show the importance of all three rings. Another interesting result was observed between the meta and para chloro substitution on the outermost rings, suggesting meta substituents are beneficial (Figure 4.5A Boxed). To test more meta-substituted sulfonamides and expand the SAR campaign beyond those commercially available, additional analogues were synthesized in collaboration with Dr. Tanpreet Kaur and Jorge Sandoval and are currently being evaluated.

A)

1 (8.68 μM)
CCG-233094
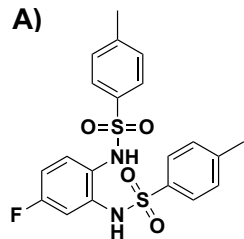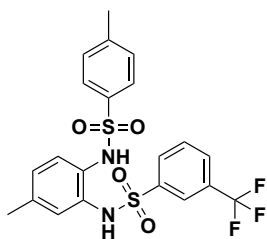
2 (10.7 μM)
CCG-234459

3 (N/A)

4 (N/A)

5 (N/A)

6 (N/A)
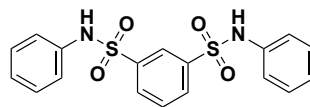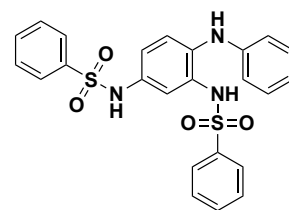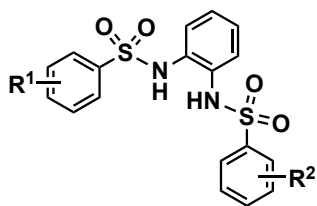
R^1	R^2	IC50 (μM)
H	H	32.1
p-F	H	14.1
p-F	p-F	28.2
p-F	p-Cl	27.2
p-F	m-Cl	11.1
p-F	p-Pr	24.2
p-F	p-OMe	17.4
p-Cl	p-Cl	27.3
o,o-di-Cl	p-F	N/A
m-Br	p-F	16.4
p-Br	p-Br	17.6
p-OMe	p-OMe	16.8
p-amide	p-amide	62.5
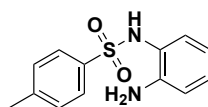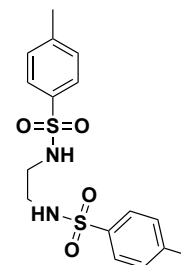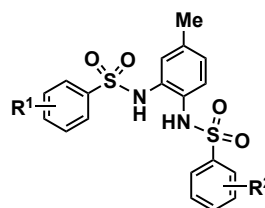
R^1	R^2	IC50 (μM)
H	H	28.6
p-F	p-F	16.4
p-Cl	p-Cl	28.3
p-Me	p-Me	14.3
p-Me	p-F	11.9

B)
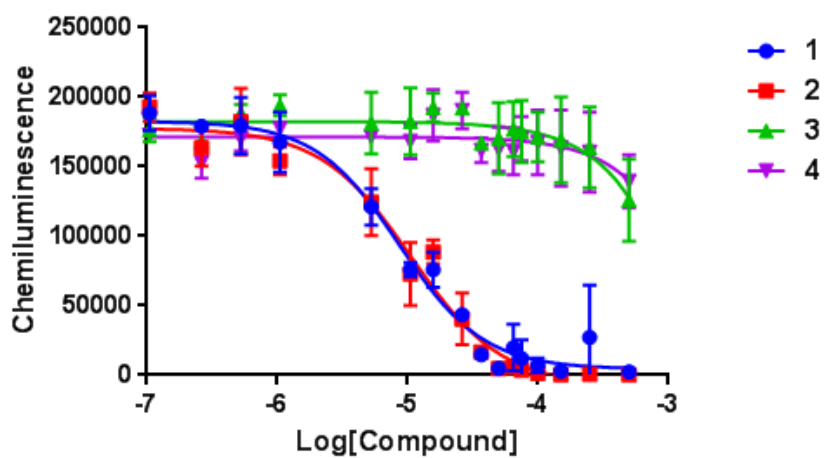
C)
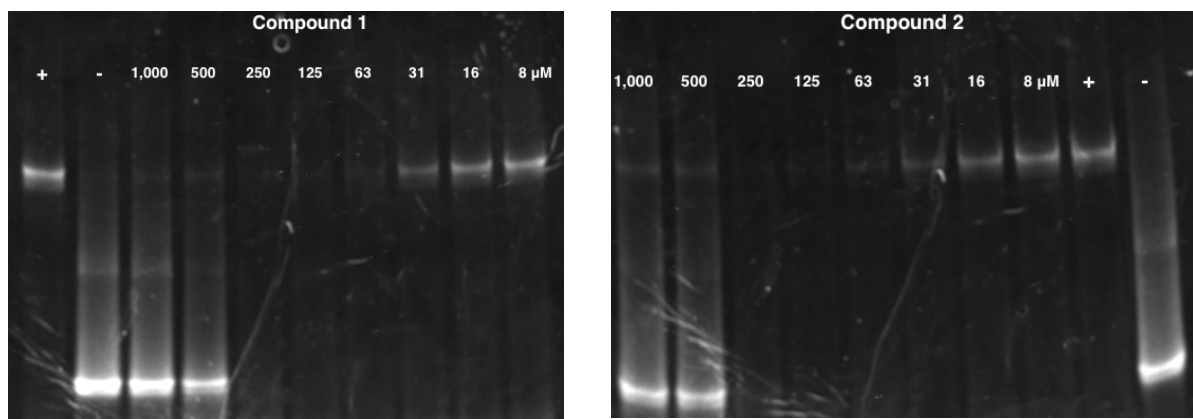
Figure 5. Sulfonamide SAR. A) Structures of sulfonamide scaffolds tested with their respective IC50 values from cat-ELCCA. Compounds with no detectable activity are listed as N/A. B) Titration curve of select compounds in cat-ELCCA. C) Electrophoretic mobility shift assay with the two top compounds. (+) represents ~~controls with Lin28 and (-) represents free RNA controls~~

Initial attempts at determining the sulfonamide mechanism-of-action focused on discerning whether the hit compounds bind the RNA or protein. In collaboration with Erin Gallagher, a biotinylated pre-let-7d and Lin28 were separately immobilized on surface plasmon resonance (SPR) chips and compound **1** was added to analyze for binding. These studies revealed that **1** had no detectable binding to the RNA and had millimolar affinity to the protein. While the high affinity to the protein was surprising, the data is likely explained by known difficulties associated with trying to determine small molecule measurements to large proteins as the difference in mass can be difficult to detect by SPR.[15] Encouraged by the idea that the sulfonamide scaffold does bind the protein, current efforts are to try other biophysical characterization techniques. Additionally, Lin28b is being expressed to determine if the compound is capable of inhibiting both homologs. Upon confirmation that the scaffold does bind Lin28, attempts will be made to obtain a co-crystal of Lin28 bound to **1**.

4.5 Conclusion

Overall, the work presented in this chapter shows the second application of cat-ELCCA for HTS and identified a hit compound **1** with a sulfonamide scaffold. Similar to other reported compounds, **1** is able to disrupt the Lin28-let7 interaction *in vitro*. With future scaffold optimizations the goal is to validate the hypothesis that manipulation of RBPs can change their biological function and potentially lead to clinically relevant outcomes.

4.6 Methods

mLin28a-halotag(HT) Cloning:

Murine Lin28a was amplified from petDuet vector generously provided by Richard Gregory (Harvard) and ligated into pFN29k using standard cloning technique with SgfI and PmeI restriction enzymes.

Primers:

5' AGTCAGCGATCGCTTCAGGCTCGGTGTCCAACCAGC

5' AAACCTTTTCAATTCTGGGCTTCTGGGAGC

mLin28a-HT Expression, Purification, and Biotinylation:

E. Coli BL21 cells transformed with the pFN29k mLin28a-HT were grown at 37°C to an OD600 of 0.6 and induced with 500μM IPTG and grown overnight at 37°C. Cells were pelleted for 15 min at 3000 x G, suspended in PBS7 (100mM Phosphate buffer pH 7, 150mM NaCl), centrifuged again for 15 min at 3000 x G and stored as pellets at -80°C until needed. Pellets were thawed at room temp before being lysed via sonication in 20 mM Imidazole pH 8, 10 mM Phosphate, 2.7 mM KCl, 137 mM NaCl, 0.1% PMSF, and 1mM DTT at 4°C. Lysates were then centrifuged at 3000 x G for 30 min at 4°C. The supernatant was then applied to Ni Resin and washed twice. First with lysis buffer then with 50/50 lysis buffer/wash buffer (10 mM Tris pH 8, 50 mM Imidazole, 500 mM NaCl, 0.1% PMSF, 1 mM DTT). Lin28-HT was eluted in 10 mM Tris pH 8, 500 mM Imidazole, 500 mM NaCl, 0.1% PMSF, 1 mM DTT before being dialyzed overnight into 20 mM Tris pH 7.8, 100 mM KCl, 0.2 mM EDTA, 10% Glycerol (v/v). Protein concentrations were

determined by Bradford using BSA standards. Biotinylation was done with 10 equivalents of Biotin-Peg-Cl ligand as described previously.[12]

RNA and HRP conjugates were prepared using previously reported methods in Chapter 3.

Pre-let7d sequence (cat-ELCCA):

The following sequence was ordered from Dharmacon:

5' N5.A.G.A.G.G.U.A.G.U.A.G.G.U.U.G.C.A.U.A.G.U.U.U.U.A.G.G.G.C.A.G.G.G.A.U.U. U.U.G.C.C.C.A.C.A.A.G.G.A.G.G.U.A.A.C.U.A.U.A.C.G.A.C.C.U.G.C.U.G.C.C.U.U.U.C. U

Lin28 cat-ELCCA:

White high binding capacity streptavidin-coated 384-well plates (Pierce 15505) were washed with 50 µL of phosphate buffer saline (100 mM phosphate, 150 mM NaCl, pH 7.0; PBS7) three times using a Biotek 405 ELX plate washer or by a multichannel pipette. 10 µL of 200 nM Lin28-HT-Biotin (in 20 mM Tris pH 7.8, 100 mM KCl, 0.2 mM EDTA, 10% Glycerol (v/v), 0.05% Tween-20) was then added by a Multidrop Combi Reagent Dispenser or multichannel pipette. Plates were then centrifuged at 1,000 RPM (223 xg) for 1 minute, sealed with plate tape and allowed to incubate overnight at 4°C. The next morning plates were washed three times with 50 µL PBS7 followed by the addition of 10µL 200nM pre-let7d-TCO in binding buffer (50 mM Tris pH 7.6, 150 mM NaCl, 5% Glycerol (v/v), 0.05% Tween-20, Fresh: 1 mM $ZnCl_2$, 10 mM Beta Mercaptoethanol. Add

DMSO to reach 5% final concentration (adjust for compounds)). Compounds were then added by a Sciclone (Caliper) liquid handler with V&P pintool or by hand pipetting. After a one hour room temperature incubation period, wells were washed with 50 µL PBS7 three times prior to the addition of 10 µL 750nM HRP-TCO in PBS7. The HRP was allowed to react for 1 hour at room temperature before being removed by three 150uL PBST (PBS7 + 0.05% Tween 20) and three 50 µL PBS7 washes. Detection was achieved by the addition of 25 µL SuperSignal West Pico (pierce). The chemiluminescent substrate was allowed to incubate at room temperature for 5 minutes before being quantified on a PHERAstar Plate reader using a LUM plus module or a Biotek Cyation3.

Electrophoretic mobility shift assay (EMSA):

SYBR Gold:

10 µL of 500 nM RNA and mLin28a-HT were mixed with DMSO or compound in binding buffer (50 mM Tris pH 7.6, 150 mM NaCl, 5% Glycerol (v/v), 0.05% Tween-20, Fresh: 1 mM ZnCl$_2$, 10 mM Beta Mercaptoethanol. Add DMSO to reach 5% final concentration (adjust for compounds)). The mixture incubated at room temperature for 30 minutes before 10µL of loading dye (binding buffer supplemented with and additional 10% glycerol and trace bromophenol blue and Xylene Cyanol for color). The samples were then loaded on 8% native TBE gels and run in TBE buffer before being stained with SYBR Gold for 5 min. Gels were then imaged on a Protein Simple gel imager using the multiflour green setting.

32P:

In an microcentrifuge tube the following were added in order: 4 µL H_2O, 1 µL 10x T4 Polynucleotide Kinase (PNK) buffer, 5 µL 100 uM RNA, 1 µL T4 PNK, 5 µL 32P Gamma ATP at 3.33 µM with approximately 8.6 mCi/ml. The labeling reaction was allowed to progress at 37°C for 30 minutes before 4 µL of 121 µM cold ATP was added and allowed to label for an additional 30 min at 37°C. PNK was inactivated by heat denaturation at 70°C for 15 min. RNA was diluted to 4nM in Binding Buffer with various concentrations of mLin28a-HT-Biotin in 4µL reaction volumes. After a 30 min. room temperature incubation the reaction was mixed with 4µL of Loading Buffer and run on a 8% native TBE gel. 32P was exposed to a phosphorimaging plate and quantified by a Typhoon Phosphoimager.

Radiolabeled Pre-let7d Sequence:

The following sequence was ordered from Dharmacon:

A.G.A.G.G.U.A.G.U.A.G.G.U.U.G.C.A.U.A.G.U.U.U.U.A.G.G.G.C.A.G.G.G.A.U.U.U.U.

G.C.C.C.A.C.A.A.G.G.A.G.G.U.A.A.C.U.A.U.A.C.G.A.C.C.U.G.C.U.G.C.C.U.U.U.C.U

4.7 References

1. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nature Publishing Group* **15,** 829–845 (2014).

2. Van Nostrand, E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., Blue, S. M., Nguyen, T. B., Surka, C., Elkins, K., Stanton, R., Rigo, F., Guttman, M. & Yeo, G. W. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nature methods* **13,** 508–514 (2016).

3. Wang, Z.-L., Li, B., Luo, Y.-X., Lin, Q., Liu, S.-R., Zhang, X.-Q., Zhou, H., Yang,

J.-H. & Qu, L.-H. Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers. *Cell Rep* **22,** 286–298 (2018).

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

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 Discov* 2472555217717944 (2017). doi:10.1177/2472555217717944

6.   Mei, H. Y., Mack, D. P., Galan, A. A., Halim, N. S., Heldsinger, A., Loo, J. A., Moreland, D. W., Sannes-Lowery, K. A., Sharmeen, L., Truong, H. N. & Czarnik, A. W. Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorganic & Medicinal Chemistry* **5,** 1173–1184 (1997).

7.   Roos, M., Pradère, 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 chemical biology* **11,** 2773–2781 (2016).

8.   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* **138,** 13630–13638 (2016).

9.   Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. *Cell* **147,** 1080–1091 (2011).

10.   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 chemical biology* **3,** 373–382 (2008).

11.   England, C. G., Luo, H. & Cai, W. HaloTag technology: a versatile platform for biomedical applications. *Bioconjugate chemistry* **26,** 975–986 (2015).

12.   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.* **19,** 763–769 (2017).

13.   Lee, H. Y., Haurwitz, R. E., Apffel, A., Zhou, K., Smart, B., Wenger, C. D., Laderman, S., Bruhn, L. & Doudna, J. A. RNA-protein analysis using a conditional CRISPR nuclease. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 5416–5421 (2013).

14. Zhang, J., Chung, T. & Oldenburg, K. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4,** 67–73 (1999).

15. Nguyen, H. H., Park, J., Kang, S. & Kim, M. Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors* **15,** 10481–10510 (2015).

Chapter 5

Real Time Homogeneous RNA-Protein Interaction Assay

The known regulatory roles of RNA-binding proteins (RBPs) in biology continue to expand rapidly due to technological advances. One such example is the improvements in cellular imaging leading to the discovery of RNA and RBP subcellular localization. These localization events have been linked to many fundamental cellular processes such as neuronal function and asymmetric cell division.[1] Current techniques that are used to rapidly identify new RNA-RBP interactions, such as CLIP (Cross Linking and Immunoprecipitation) and proteomics, fail to account for cellular localization or only provide a snapshot of interactions. This highlights the need for the development of new technologies that allow for the high-throughput detection of these dynamic RBP-RNA interactions in live cells that can function on a subcellular level.

One technique used to study protein-protein interactions (PPIs) in live cells is the protein complementation assay (PCA).[2] A PCA works by taking a detection protein, like luciferase or green fluorescent protein, and splitting it into two parts. These parts are designed to have a low affinity for each other and minimal activity when separated but regain activity when driven together by an external force, such as a known PPI. For a PPI PCA, the PCA protein fragments are expressed as fusion proteins to known PPI partners allowing for signal to be produced when the PPI partners interact naturally. This system

has been implemented for many PPIs and, highlighting the dynamic detection, has been used to study changes to PPIs upon external stimuli.[3] Additionally, PCAs have been used in a live cell high-throughput screening (HTS) settings to identify novel PPI partners.[4] However, this system has never been reported for studying RNA-protein interactions (RPIs) due to the inability to generate RNA-protein fusions.

5.1 RNA Binding Protein Complementation Assay In Vitro

Initial designs and proof-of-concept experiments for a RPI PCA focused on using the Inverse Electron Demand Diels Alder (IEDDA) reaction to generate a RNA-protein conjugate due to the previous success with cat-ELCCA (Figure 5.1).[5] While there are many PCA systems, few have been used in an in vitro setting.[6,7] One of these systems, a split firefly luciferase, displayed instability when expressed recombinantly and needed to be used fresh.[8] This instability would be a major limitation for the design of a RPI-PCA. Alternatively, Promega was able to develop a PCA, termed NanoBiT, for monitoring PPIs that was ideal for both *in vitro* and *in celluo* applications.[9] The NanoBiT system was generated by splitting NanoLuc luciferase into two fragments, LgBiT and SmBiT. To apply the NanoBiT assay to RPIs, a Halotag(HT) SmBiT fusion protein was generated to serve as a conjugation site for RNA. The LgBiT could then be fused to a RBP-of-interest to complete the system. As proof-of-concept, the Lin28-Let-7 system was chosen because it is the most well-characterized pre-miRNA-protein interaction (Figure 5.1). Both of these fusion proteins were expressed in *E. coli* using standard techniques and found to be stable, including the HT-SmBiT when conjugated with a methyltetrazine (mTet).
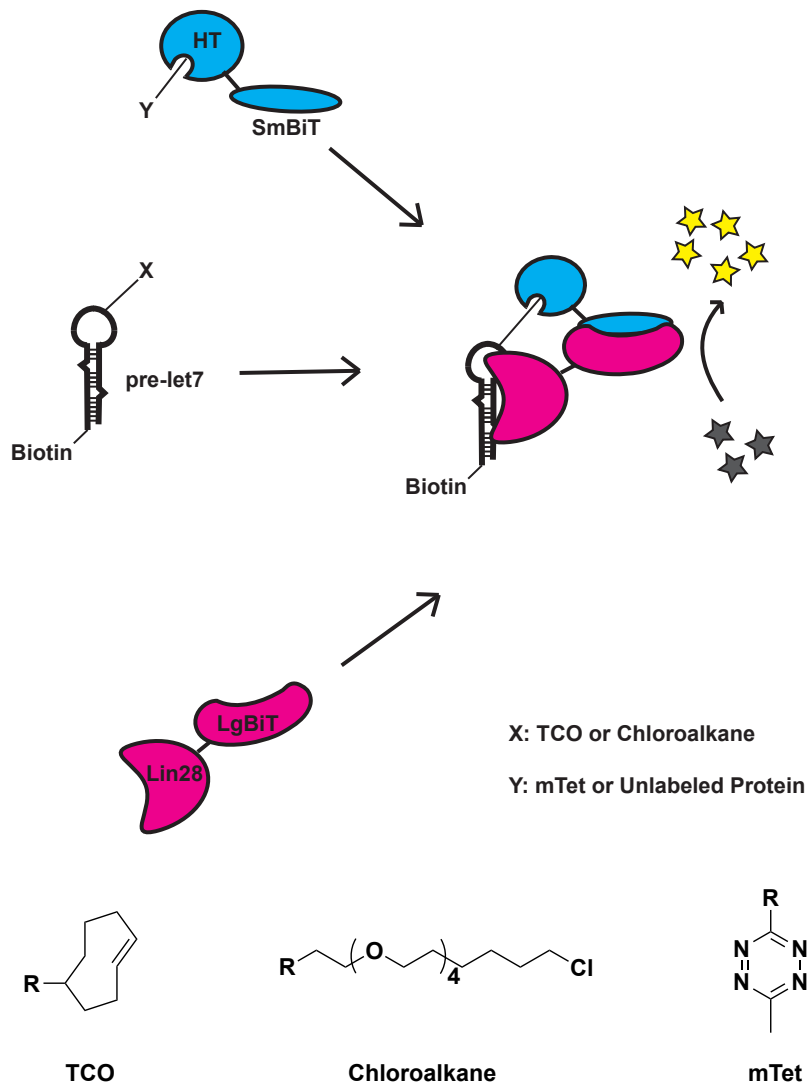
Figure 5.1. RBP-PCA Scheme. Schematic for in vitro RNA-Protein interactions protein complementation assay.

After generating all of the necessary components, several experiments were carried out to test the RPI PCA. The first was mixing all of the required components, HT-SmBiT-mTet and Lin28a-LgBiT with varying concentrations of pre-let-7d-TCO. Initial attempts to detect binding failed with the commercial live cell reagent due to high background noise; however, upon switching to the lytic detection reagent a change in

signal relative to the concentration of RNA was observed (Figure 5.2A). This change was attributed to the presence of detergents and other reagents that disrupt low affinity interactions that are present in lytic buffers preventing the non-specific association of SmBiT and LgBiT. To confirm that the increase in signal was due to the Lin28-pre-let-7d interaction, the signal was competed with unlabeled Lin28a and a corresponding decrease in intensity was observed (Figure 5.2B). Intrigued by other potential applications for a RBP PCA, direct RNA to HT coupling was attempted *in vitro* using a chloroalkane-labeled RNA. This orthogonal coupling strategy was successful, as only pre-let-7d labeled with a chloroalkane produced signal (Figure 5.2C). However, it should be noted that the overall signal intensity and signal-to-background was less than the click chemistry variant. This decrease is likely the result of stoichiometric addition of chloro-labeled RNA and SmBiT-HT, as compared to the excess of chloro-mTet used to generate the SmBiT-HT-mTet. Finally, to assess the assay's ability to respond to real time changes, RNase A was added to degrade the pre-let-7d substrate. As shown in Figure 5.2D, an immediate decrease in signal was observed upon the addition of RNase A. Taken together these experiments show that protein complementation assays can be applied to RPIs *in vitro*.
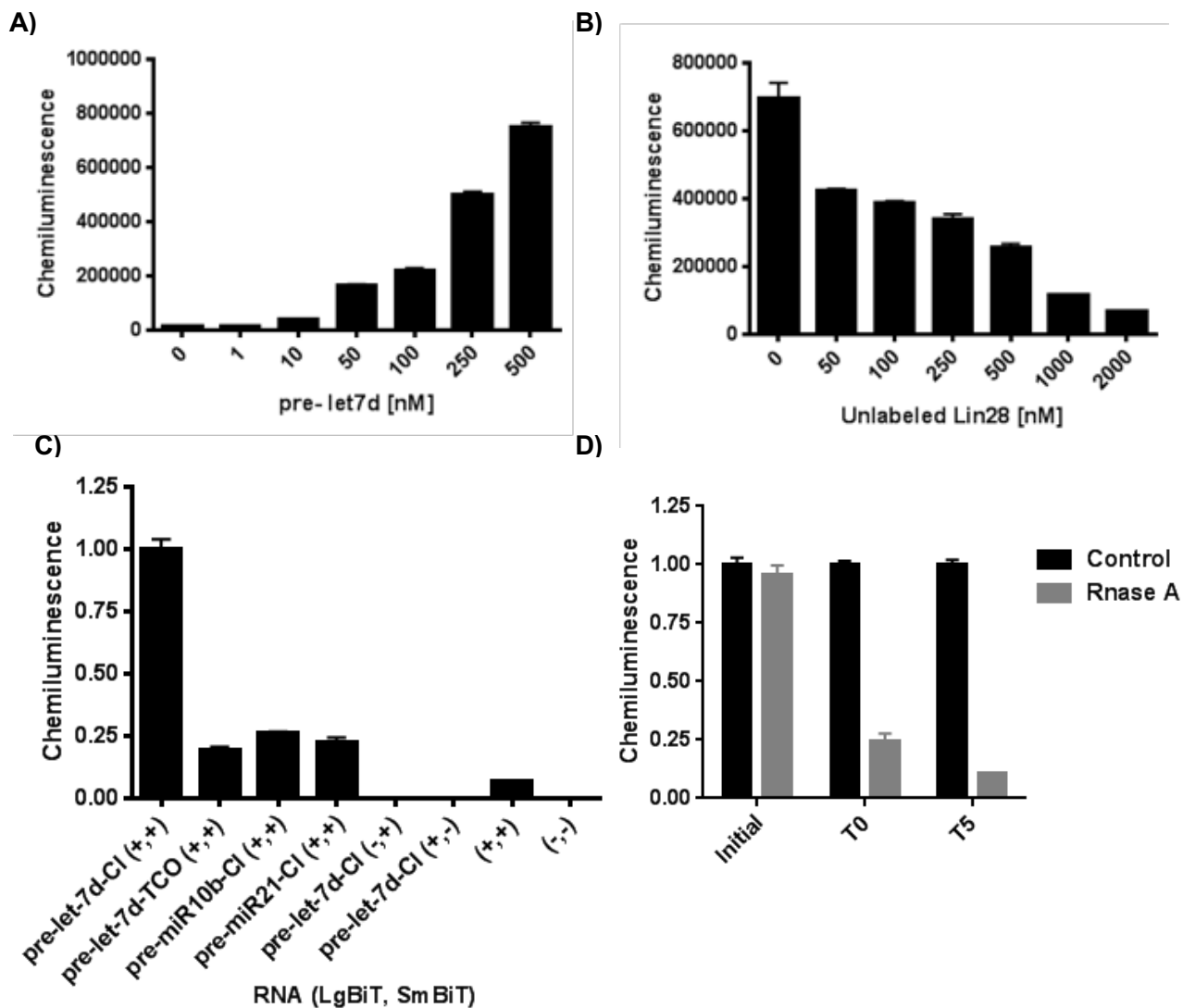
Figure 5.2. In Vitro PCA Proof-of-Concept. A) Titration of labeled pre-let7d-TCO in RPI-PCA. B) Competition of 500nM pre-let7d-TCO and Lin28-LgBiT with Lin28-HT-Biotin. C) RPI-PCA with chloroalkane RNA. D) Time dependence of click RPI-PCA with 1ug RNase A at initial read, T0 = signal after addition of RNase A, and T5 = 5 min. post RNase addition.

5.2 Monitoring RNA Protein Interactions in Live Cells

Having generated a functional PCA for RPIs in vitro, the next goal was to apply this new technology to monitory RPIs in live cells. This cell-based assay was designed as seen in Figure 5.3A. Central to the initial design was the generation of a stable cell line expressing SmBiT-HT to simplify the number of components needed to add for each experiment. SmBiT-HT was the ideal candidate for genomic incorporation, as it would be universal for all potential cytoplasmic RBP partners. HEK293-FlpIn cells were chosen as a model system to generate this stable cell line because it contains a site for specific genomic recombination, ensuring a monoclonal cell line, and is commercially available. As can be seen in Figure 5.3B, the FlpIn cell line was able to stably express the SmBiT-HT protein and transiently express Lin28a-LgBiT. Initial attempts to transfect 100 ng of Lin28a-LgBit plasmid per 96-well containing 10,000 cells failed to produce the desired signal. However, upon reduction in the amount of DNA to 1.25 ng, a clear signal separation between pre-let-7 and control RNAs was observed. The specificity of this new assay was then evaluated by transfecting different pre-miRNAs that either lacked the required chloroalkane tag or are not reported to bind Lin28. As can be seen from Figure 5.3C, only the correct combination of Lin28-LgBiT, SmBiT-HT, and pre-let-7d-Cl produced the desired signal.
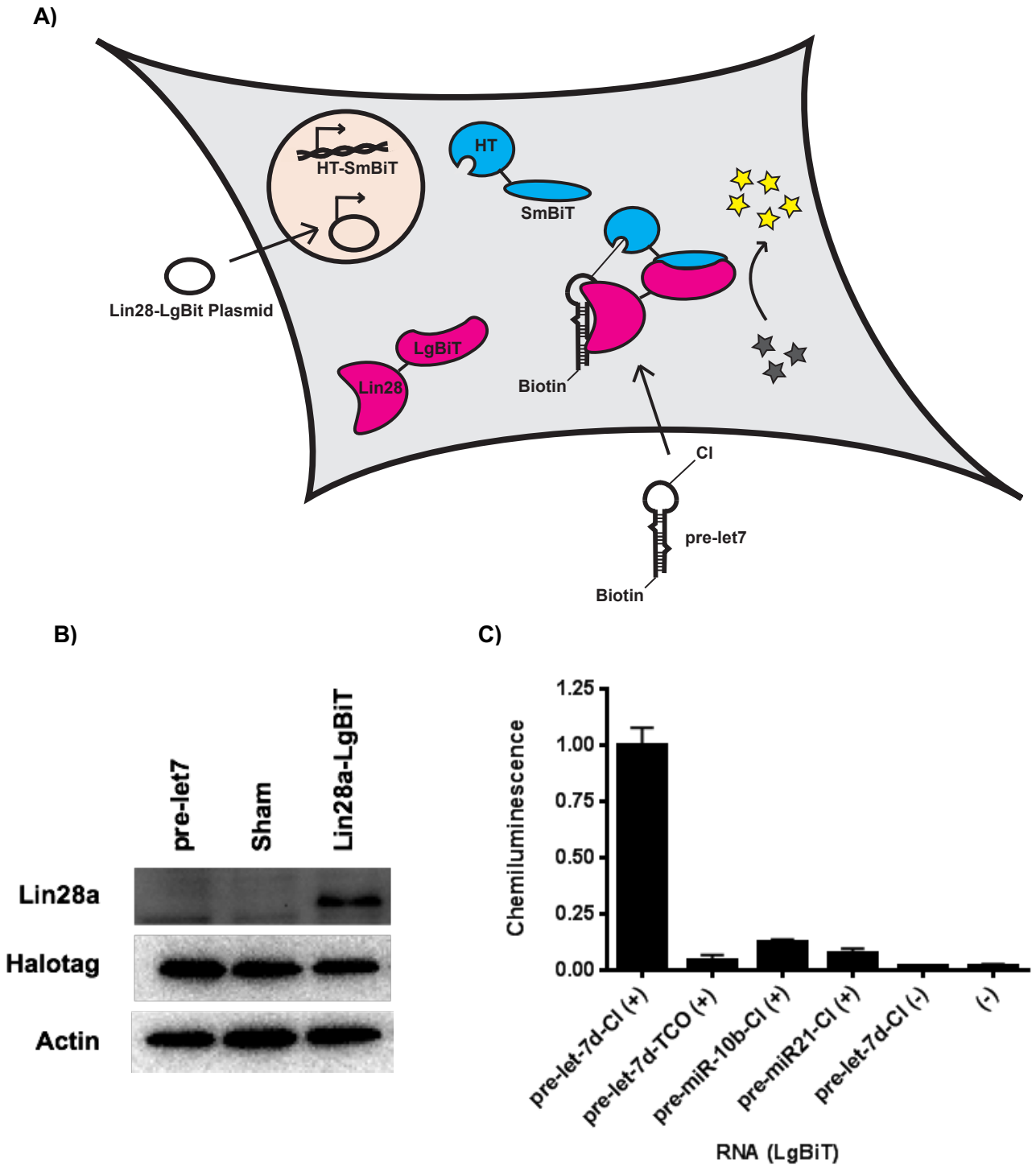
Figure 5.3. Lin28 Cell Based PCA. A) Schematic of cell based RPI-PCA. B) Western blots showing expression of Halotag-SmBiT and Lin28a-LgBiT. C) Cell based RPI-PCA with select RNAs and no Lin28 transfection controls

5.3 Identification of pre-miRNA Binding Proteins

Ever since the discovery of the let-7-Lin28 system, many research groups have focused on the identification of similar systems. Backed by evidence of conserved loop regions in pri- and pre-miRNAs, these identification efforts have typically relied on protein pull-down experiments using RNA as a bait followed by proteomics.[10] In collaboration with Emily Sherman, a pipeline for the discovery of new RBPs using the biotinylated pre-miRNA probe, designed for both Dicer cat-ELCCA and *in vitro* PCA, was generated (Figure 5.4A). Using pre-let-7d and pre-miR-21 RNAs, a pull-down experiment was performed and analyzed by Western blot prior to submission for proteomic analysis. Both the Western blot and proteomic data showed significant enrichment of Lin28 in the pre-let-7d sample over pre-miR-21 as expected (Figure 5.4B and C). Upon examination of the proteomic data, another pre-let-7-specific RBP was identified, GRSF1 (G-rich sequence factor 1) (Figure 5.4C). This interaction has been reported by others as well.[11] While GRSF1 has been shown to pull down with pre-let-7, there are no reports demonstrating this interaction *in celluo*.

A)

B)

	Input		FT		Resin	
	7	21	7	21	7	21
Lin28a						

C)

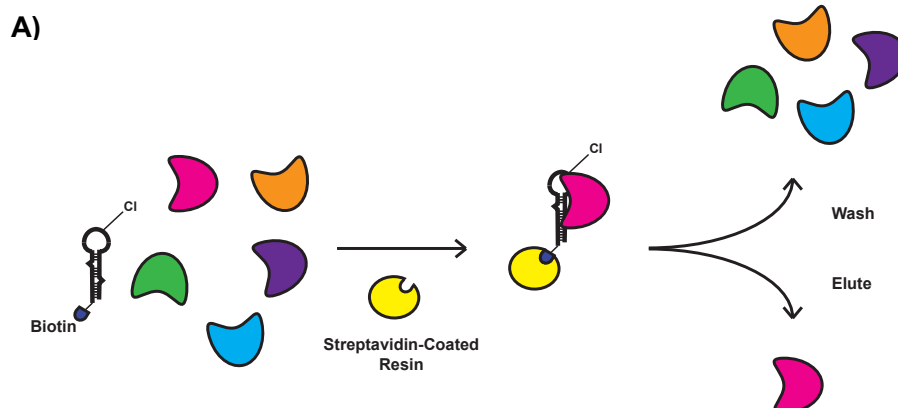| Protein	PSM (pre-let7d)	PSM (pre-miR-21)
LIN28b	28	11
GRSF1	138	2

Figure 5.4. pre-miRNA RBP Proteomics. A) Workflow for pull-down of pre-miRNA binding proteins. B) Western blot from a pull-down experiment showing enrichment of Lin28 when pre-let-7d was used as bait over control pre-miR-21. C) Peptide spectrum matches (PSMs) of Lin28b and GRSF1 using two different RNA baits.

5.4 GRSF1 Binds Pre-let-7d in the Mitochondria

GRSF1 is a RBP that has been shown to regulate mitochondrial gene expression.[12] This gene regulation is thought to be due to GRSF1's ability to bring non-coding RNAs from the cytoplasm into the mitochondria.[13] Coupling GRSF1's ability to shuttle RNA to the mitochondrial matrix with reports of mitochondrial miRNAs led to the hypothesis that GRSF1 shuttles the let-7 family into the mitochondrial matrix.[14] Central to this hypothesis is that both GRSF1 and pre-let-7 associate in the cell. To probe this interaction, the cell-based RPI PCA assay was used. Unfortunately, after transfecting a construct containing GRSF1-LgBiT into the FlpIn HEK cells, no significant signal was observed. Undeterred by this, it was reasoned that SmBiT-HT was not present in the mitochondrial matrix where GRSF1 is known to localize. To ensure proper colocalization of the two BiTs, the N-terminus of GRSF1 that is predicted to cause mitochondrial localization was cloned onto the SmBiT-HT. Additionally, instead of generating a stable cell line for every compartment of the cell, the cell-based RBP PCA was tested in traditional HEK293 cells by transfecting all three components at once (SmBiT-HT, RBP-LgBiT, RNA). Using the Lin28 system as a control, this triple transfection was successful (Figure 5.5A). Upon transfection of GRSF1-LgBiT, the cell-based PCA assay performed

as hypothesized with robust signal only observed when both BiTs were localized to the mitochondria (Figure 5.5B). Importantly, this signal intensity was dependent on pre-let-7d, as both no RNA and a control RNA, pre-miR-34, displayed lower activity levels. This data provides the first direct evidence that GRSF1 and pre-let-7d interact in the mitochondria. Future work will be needed to evaluate the changes of mitochondrial miRNA levels upon GRSF1 knock down to continue to test the hypothesis that GRSF1 acts as a miRNA shuttle.

A) **mLin28a-LgBiT**
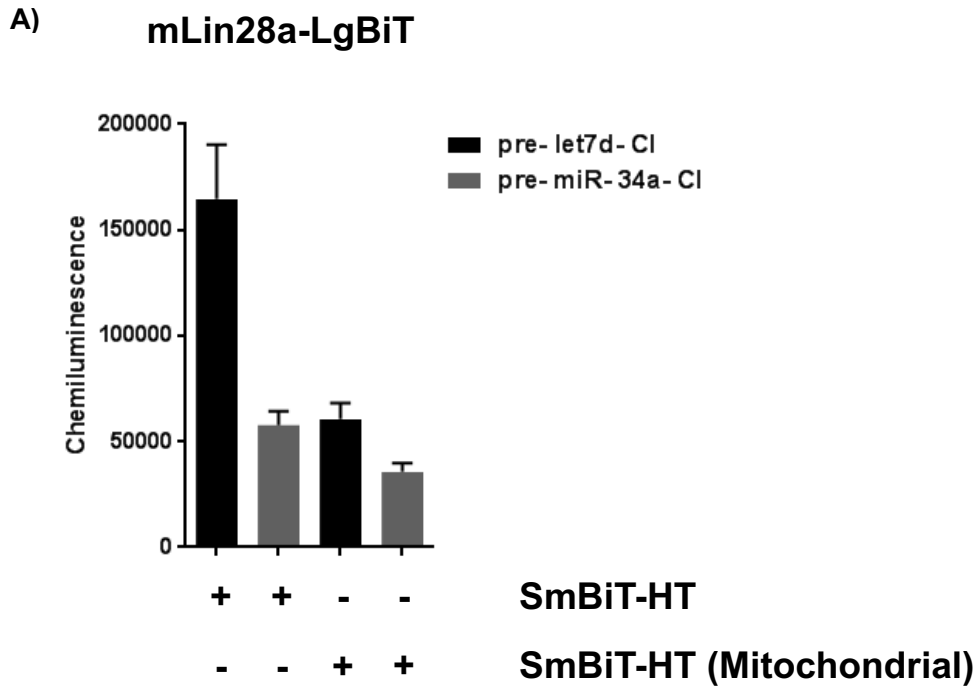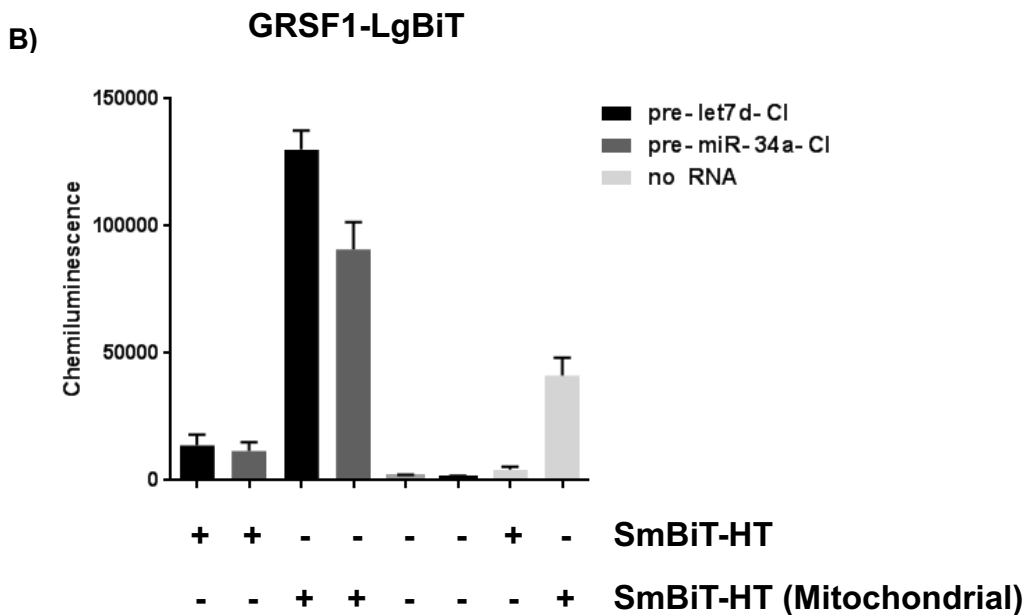
Figure 5.5. GRSF1 Cell Based PCA. A) Lin28-pre-let7d control experiment with HEK293T cells B) Signal enrichment between GRSF1 and pre-let7d over control pre-miR-34a

5.5 Conclusion

The work presented in this chapter continues to highlight the potential uses of RNA bioconjugation for both *in vitro* and *in celluo* applications. Using both click chemistry and Halotag coupling, the first *in vitro* RPI-PCA was developed for Lin28-pre-let-7. This in vitro PCA assay served as a proof-of-concept to generate a method to monitor RPIs in live cells. The value of this *in celluo* assay was highlighted by the confirmation of the GRSF1-pre-let7d interaction localized to the mitochondria.

5.6 Methods

Constructs:

All protocols were based on traditional restriction enzyme cloning protocols supplied from New England Biolabs.

In vitro Constructs:

LgBiT and SmBiT were cloned from pFN33k and pFN35k into pFN29k and pFC30k, respectively, to generate halotag(HT) fusion proteins using primer pairs **1** and **2**. The halotag was then removed from the pFN29k LgBiT construct using primer pair **3** and replaced with mLin28a using primer pair **4** followed by the addition of primer pair **5** to add a 6x his tag to generate the final construct.

Cell-Based Constructs:

mLin28a-LgBiT and SmBiT-HT were cloned from the in vitro constructs into either pcDNA3 or pcDNA5/FRT for either transfections or stable cell line generations using primer pairs **6**, **7**, and **8**. GRSF1 gene was synthezied to contain KpnI and SgfI cut sites due to challenges cloning the GC rich region and cloned into pcDNA3 containing LgBiT. The N-terminus of GRSF1 was cloned using primer pair **9** to the N-terminus of SmBiT-HT in pcDNA3.

Table 5.1. Primers for Construct Generation

Pair	Primer	Seq	Comments
1	LgBiT29k F	gagcgataacgcgatcgatggtcttcacactcgaagatttcgttgg	From pFN33k to pFN29k
1	LgBiT29k R	cgaattcgtttaaacactgttgatggttactcggaacagc	
2	SmBiT30k F	cgccatggtgaccggctaccggctgttcgaggagattctcc	From pFN25k to pFC30k
2	SmBiT30k R	tcgaggagaatctcctcgaacagccggtagccggtcaccatggcgat	
3	pfn29khindiii F	ggtattttaactttactaaggagaagcttcatcatgaaacatcatcacc	pFN29k mutagenesis to add N term HindIII site
3	pfn29khindiii R	ggtgatgatgtttcatgatgaagcttctccttagtaaagttaaaatacc	
4	Lin28 into pFN29k F	ctaaggagaagcttatgggctcggtgtccaaccagc	From pFN29k Halotag (Chapter 4) into LgBiT vector lacking HT
4	Lin28 into pFN29k R	ggtcacggcggatcgcattctgggcttctgggagc	
5	His Tag	agctgatgggtcaccatcaccatcaccata	Annealed oligos to add his tag in HindIII site
5	His Tag	agcttatggtgatggtgatggtgacccatc	
6	SmHT30 SEA F	gtacggtaccgccaccatggtgaccggctaccgg	SmBiT-HT (30k vector) to pcDNA5/FRT: PCR out the SmBit-HT fusion, add KpnI, Kozak sequence, and NotI (used for stable cell line generation)
6	SmHT30 SEA R	cagtgcggccgctcactattagtggtgatggtgatgatg	
7	SmHT30Blp-Rev	gactgcggccgctcattagctcagcccaccggaaatctccagagtag	Smbit30k rev. remove stop from gene, add Blpl then stop codons then notl (for transient transfection in HEK cells)
8	LinLG29 SEA F	gtacggtaccgccaccatgggctcggtgtccaacc	Lin28a-LgBit (29k vector) to pcDNA3: PCR out the lin28a-LgBit fusion, add KpnI, Kozak sequence, and NotI
8	LinLG29 SEA R	cagtgcggccgctcattaaacactgttgatggttactcg	
9	N-term GRSF1 F	gtacggtaccgccaccatggccggcacgcgctg	PCR N-Term of GRSF1 from purchased gene into pcDNA3 SmHT30 (Nterm via KpnI)
9	N-term GRSF1 R	agtcggtaccgtagctgcgcgtcgggac	

Protein Expression:

E. coli BL21 were transfected with the mLin28a-LgBiT (pFN29K) and SmBiT-HT (pFN30k). 1 L cultures were then grown to an OD600 of 0.6 and induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside and allowed to grow overnight at 37°C. The purification protocol of mLin28a-HT from chapter 5 was used for both protein fusions. mTetrazine halotag ligand was added to purified protein at 20 equivalents and allowed to react overnight at 4°C before excess was buffer exchanged away.

mTetrazine halotag ligand was synthesized using the reported protocol[15]

In Vitro PCA:

500 nM SmBiT-HT (+/- mTet), Lin28a-LgBiT, and RNA (TCO/Cl) were mixed 10 µL Binding Buffer (50 mM Tris pH 7.6, 150 mM NaCl, 5% Glycerol (v/v), 0.05% Tween-20, Fresh: 1mM $ZnCl_2$, 10mM Beta Mercaptoethanol. Add DMSO to reach 5% final concentration (adjust for compounds)). Following a 1-hour incubation at room temperature, 5 µL of NanoGlo reagent from Promega was added (1:50 dilution of reagent to buffer). Chemiluminescence was detected after 5 minutes using a Biotek Cyation 3.

Generation of HEK293 FlpIn stable cell lines:

HEK293 FlpIn cells were purchased from Fisher Scientific. Generation of the SmBiT-HT cell line was achieved by following the suggested protocol from the vender.

Cell Based Protein Complementation Assay:

HEK293 and HEK FlpIn(SmBiT-HT) cells were cultured using standard cell culture techniques in DMEM supplemented with 10% fetal bovine serum, 2 mM Glutamine, and 1% Pen-Strep (w/v). 400 µL of 100,000 cells per mL was prepared for reverse transfection with 50 µL optimem containing 1.2 µL RNAi Max Lipofectamine (Life Technologies), 1.25 µL RBP-LgBiT (4ng/µL) pcDNA3 plasmid, 1.25 µL SmBiT-HT (4 ng/µL) pcDNA3 plasmid (not applicable for FlpIn cells), and 0.3 µL 50 µM RNA. 100 µL was then added to each well of a 96-well white tissue culture microtiter plate and allowed to incubate for 24 hours at 37°C with 5% CO_2. Media was then aspirated and replaced with 100 µL room temperature Opti-MEM. 25 µL of NanoGlo Live Cell Reagent (Promega) was then added to each well (1:20 dilution of reagent to buffer), allowed to incubate for at least 5 minutes before detecting total chemiluminescence. Signal was stable for roughly 1 hour.

5.7 References

1. Martin, K. C. & Ephrussi, A. mRNA localization: gene expression in the spatial dimension. *Cell* **136,** 719–730 (2009).

2. Morell, M., Ventura, S. & Avilés, F. X. Protein complementation assays: approaches for the in vivo analysis of protein interactions. *FEBS Letters* **583,** 1684–1691 (2009).

3. Luker, K. E., Smith, M. C. P., Luker, G. D., Gammon, S. T., Piwnica-Worms, H. & Piwnica-Worms, D. Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals. *Proceedings of the National Academy of Sciences* **101,** 12288–12293 (2004).

4. MacDonald, M. L., Lamerdin, J., Owens, S., Keon, B. H., Bilter, G. K., Shang, Z., Huang, Z., Yu, H., Dias, J., Minami, T., Michnick, S. W. & Westwick, J. K. Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nature Chemical Biology* **2,** 329–337 (2006).

5. Lorenz, D. A. & Garner, A. L. A click chemistry-based microRNA maturation assay

optimized for high-throughput screening. *Chem. Commun.* **52,** 8267–8270 (2016).

6.   Hashimoto, J., Watanabe, T., Seki, T., Karasawa, S., Izumikawa, M., Seki, T., Iemura, S.-I., Natsume, T., Nomura, N., Goshima, N., Miyawaki, A., Takagi, M. & Shin-Ya, K. Novel in vitro protein fragment complementation assay applicable to high-throughput screening in a 1536-well format. *Journal of Biomolecular Screening* **14,** 970–979 (2009).

7.   Ohmuro-Matsuyama, Y., Chung, C.-I. & Ueda, H. Demonstration of protein-fragment complementation assay using purified firefly luciferase fragments. *BMC Biotechnol.* **13,** 31 (2013).

8.   Ohmuro-Matsuyama, Y., Hara, Y. & Ueda, H. Improved Protein–Protein Interaction Assay FlimPIA by the Entrapment of Luciferase Conformation. *Analytical chemistry* **86,** 2013–2018 (2014).

9.   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* **11,** 400–408 (2016).

10.   Michlewski, G., Guil, S., Semple, C. A. & Cáceres, J. F. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Molecular Cell* **32,** 383–393 (2008).

11.   Newman, M. A., Thomson, J. M. & Hammond, S. M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14,** 1539–1549 (2008).

12.   Antonicka, H., Sasarman, F., Nishimura, T., Paupe, V. & Shoubridge, E. A. The mitochondrial RNA-binding protein GRSF1 localizes to RNA granules and is required for posttranscriptional mitochondrial gene expression. *Cell Metabolism* **17,** 386–398 (2013).

13.   Noh, J. H., Kim, K. M., Abdelmohsen, K., Yoon, J.-H., Panda, A. C., Munk, R., Kim, J., Curtis, J., Moad, C. A., Wohler, C. M., Indig, F. E., de Paula, W., Dudekula, D. B., De, S., Piao, Y., Yang, X., Martindale, J. L., de Cabo, R. & Gorospe, M. HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. *Genes & Development* **30,** 1224–1239 (2016).

14.   Barrey, E., Saint-Auret, G., Bonnamy, B., Damas, D., Boyer, O. & Gidrol, X. Pre-microRNA and mature microRNA in human mitochondria. *PLoS ONE* **6,** e20220 (2011).

15.   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.* **19,** 763–769 (2017).

Chapter 6

Conclusions and Future Directions

Bioconjugation techniques have been used by chemical biologist to study a myriad of targets; however, one area of biology for which this toolkit has only been minimally applied is miRNAs. Despite miRNAs playing a role in nearly every aspect of biology, there has been a gap in our ability to manipulate these important micromanagers.[1,2] The work presented in Chapters 3, 4, and 5 highlight the application of chemical biology to the study and targeting of miRNAs. The development of two separate cat-ELCCA assays discussed in Chapters 3 and 4 for Dicer-dependent miRNA maturation and for the Lin28-pre-let-7d RNA-protein interaction resulted in the identification of new scaffolds to manipulate miRNA biology. Additionally, the modularity designed into each cat-ELCCA assay will enable quick adaptation to other pre-miRNAs or RNA-protein targets-of-interest. Chapter 5 described the use of similar bioconjugation techniques to facilitate the process of validating new RNA-protein interactions (RPIs) through a combination of proteomics and a new RPI protein complementation assay (PCA). When these technologies are combined, they create a pathway to discover new biology for a miRNA-of-interest and the tools to manipulate that biology (Figure 6.1).
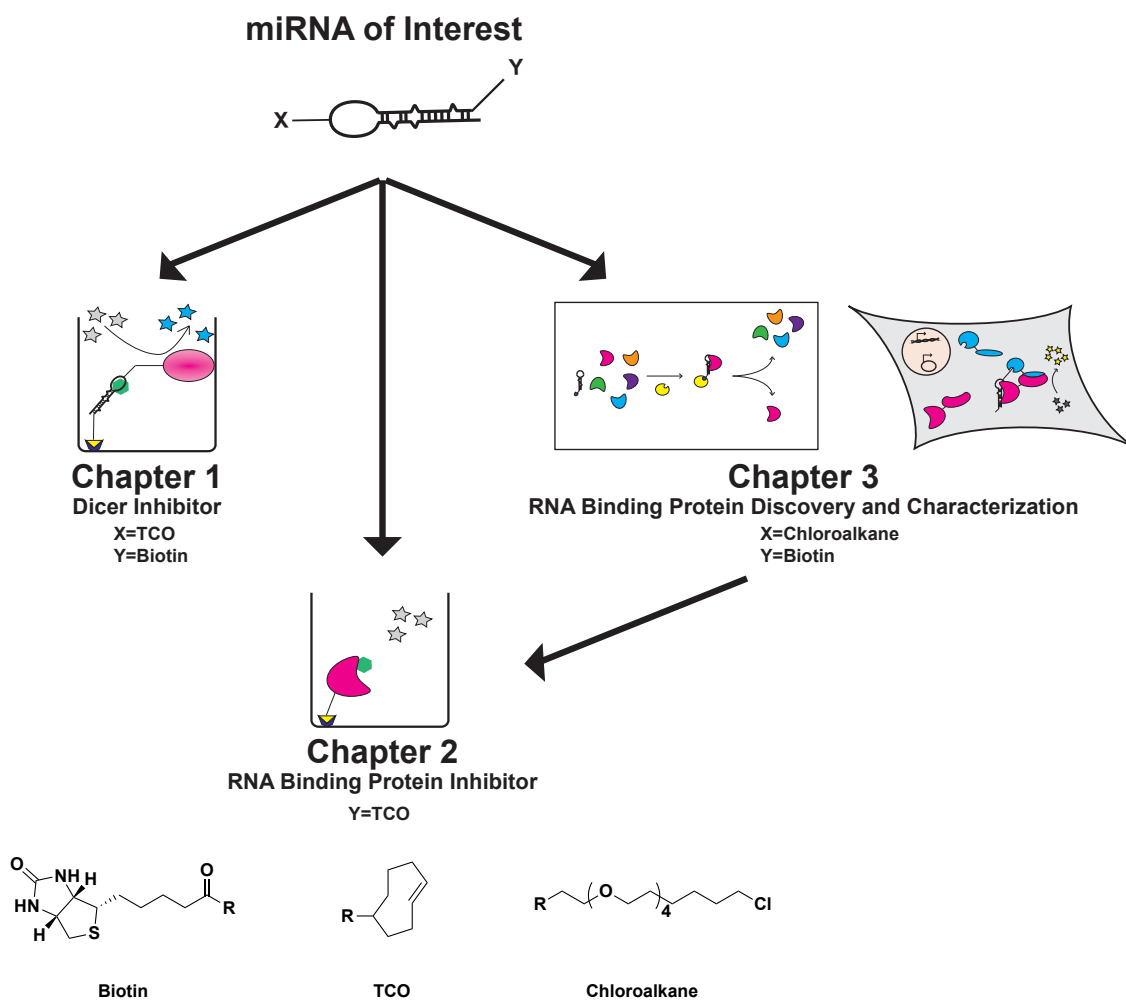
Figure 6.1. miRNA Technology Pipeline. Pipeline for the discovery of miRNA pre-miRNA manipulators using technology outlined in previous chapters.

6.1 The Future of Dicer Inhibitors

The information discussed in Chapter 1 highlights how proper control of miRNA expression is important to maintain cellular health and homeostasis. As mentioned in Chapter 2, the development of RNA-targeted oligonucleotide and small molecule drugs has been hindered by either tissue distribution or target selectivity, respectively. This conclusion is supported by the results in Chapter 3, which demonstrated that what we

currently consider drug-like compounds and scaffolds for targeting RNA lack the properties necessary for selective pre-miRNA inhibition. Using RNA regulation found in biology as an inspiration, it is surely possible to find a selective compound from commercial compound libraries, but it should probably be viewed as a needle-in-a-haystack approach. Therefore, the discovery of new scaffolds that bind RNA, and in particular pre-miRNAs, will be paramount to understanding small molecule-RNA interactions. Once a strong understanding of these interactions is obtained, the knowledge could then be applied to design molecules with selectivity. Chapter 3 described the establishment of a platform for which to screen natural product extracts as a viable strategy to find new RNA-targeted chemical space. Future work elucidating the mechanisms by which these compounds bind to RNA will prove to be valuable to the field of RNA-targeted small molecules. Furthermore, expanding the source of natural products beyond actinobacteria could provide even more unique scaffolds and potentially reveal novel mechanisms for RNA regulation in Nature.

6.2 Manipulating RNA Binding Proteins

As mentioned previously, the successful manipulation of select RNAs with small molecules has remain limited. Chapters 2 and 4 provided evidence that targeting known RPIs, such as Lin28-let-7, with small molecules can result in RNA modulation. Having identified an active sulfonamide scaffold using cat-ELCCA, future work is underway to fully characterize the interaction to improve the potency and expand our knowledge of targeting RNA-binding proteins (RBPs). However, it will be critical to show both *in celluo* and *in vivo* activity, as this has yet to be demonstrated for Lin28-let-7 inhibitors. Should

such activity be demonstrated, an entire new strategy for therapeutic manipulation will emerge, as RBPs play a role in essentially every biological pathway.

6.3 Characterization of RNA-Binding Proteins

With many RBPs yet to be characterized, the field of RNA research has countless questions left unanswered. To help address many of these questions, a new cell based RPI assay was described in Chapter 5 and was used to show the mitochondrial co-localization of pre-let-7d and G-rich sequence factor 1 (GRSF1). Future work for this RPI-PCA should focus on optimizing signal-to-background. These optimizations could be achieved through a variety of modifications including using a more optimal chloroalkane linker and adjusting transfection reagents.[3] Additionally, with the recent incorporation of unnatural nucleotides in cells, it is theoretically possible to incorporate the necessary modifications under endogenous control.[4,5] As the technology continues to be developed, the RPI-PCA could be used in variety of ways including a RNA interactome screen by fusing the known RBPs to LgBiT, as well as profiling changes during development in real time.

6.4 Concluding Remarks

Small molecule drug discovery is on the verge of a renaissance that is set up by our understanding of RNA biology. The tipping point is being catalyzed by development of tools such as the ones in this thesis that enable the rapid discovery and adaptation to many RNA targets. Once the first breakthrough is achieved by a small molecule RNA modulating drug, a new huge repertoire of clinically relevant targets awaits. Using the

105

information gained along the way will prove invaluable in enabling the future of small molecule RNA therapeutics.

6.5 References

1. Connelly, C. M., Moon, M. H. & Schneekloth, J. S. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chemical Biology* **23,** 1077–1090 (2016).

2. Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature reviews. Drug discovery* **16,** 203–222 (2017).

3. Friedman Ohana, R., Kirkland, T. A., Woodroofe, C. C., Levin, S., Uyeda, H. T., Otto, P., Hurst, R., Robers, M. B., Zimmerman, K., Encell, L. P. & Wood, K. V. Deciphering the Cellular Targets of Bioactive Compounds Using a Chloroalkane Capture Tag. *ACS chemical biology* **10,** 2316–2324 (2015).

4. Feldman, A. W., Dien, V. T. & Romesberg, F. E. Chemical Stabilization of Unnatural Nucleotide Triphosphates for the in Vivo Expansion of the Genetic Alphabet. *Journal of the American Chemical Society* **139,** 2464–2467 (2017).

5. Zhang, Y., Lamb, B. M., Feldman, A. W., Zhou, A. X., Lavergne, T., Li, L. & Romesberg, F. E. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proceedings of the National Academy of Sciences* **114,** 1317–1322 (2017).