UNIVERSITY OF CALIFORNIA

Los Angeles

Primary microRNA processing in humans:

biology, biotechnology and disease

A dissertation submitted in partial satisfaction of the

Requirements for the degree Doctor of Philosophy

In Biological Chemistry

by

Jen Elana Quick-Cleveland

2017

ABSTRACT OF THE DISSERTATION

Primary microRNA processing in humans;

biology, disease, and biotechnology

by

Jen Elana Quick-Cleveland

Doctor of Philosophy in Biological Chemistry

University of California, Los Angeles, 2017

Professor Feng Guo, Chair

MicroRNAs (miRNAs) are a class of non-coding RNAs that tune gene expression by negatively regulating at least 60% of protein-coding genes. In animals, they are required for development and cell physiology. Due to their role in controlling gene expression, dysregulation of miRNA biogenesis can cause genetic disorders, immunity deficits, neurological problems, and cancers. miRNA genes are transcribed into primary miRNA (pri-miRNAs), which undergo a multi-step maturation process. The first step occurs in the nucleus, where the pri-miRNA is cleaved by the Microprocessor Complex (MC). The MC contains the ribonuclease Drosha and an RNA-binding hemoprotein DGCR8. The MC specifically recognizes the characteristic pri-miRNA hairpin to initiate miRNA maturation. This thesis centers on the nuclear step of miRNA biogenesis. My research has two main areas of focus; 1) dissecting the molecular and structural determinants that govern pri-miRNA processing, and 2) investigating the regulation of pri-miRNA processing in normal physiology and diseases.

The MC identifies pri-miRNA substrates from a myriad of other RNAs. DGCR8 plays an important role in pri-miRNA recognition, but the mechanism was unknown. DGRC8 contains two double-stranded RNA-binding domains (dsRBDs), but these domains do not provide specificity. It has been shown that “junctions” between single-stranded and double-stranded regions in pri-miRNAs are important features that define MC substrates. However, the protein moiety that recognizes pri-miRNA junctions had not been identified. We discovered that DGCR8 contains an RNA-binding heme domain (Rhed) that directly and specifically binds pri-miRNA junctions. Further, we showed that the Rhed and its RNA-binding surface are required for efficient pri-miRNA processing.

Previous studies of our lab showed that the pri-miRNA processing activity of DGCR8 specifically requires heme in its Fe(III) redox state. However, it is unknown how much Fe(III) heme is available in cells and how much is required to support pri-miRNA processing. During our investigation of the Rhed-pri-miRNA interface, we performed mutagenesis on DGCR8 that led to identification of single mutations with reduced heme affinity but nearly full processing activity in cells. This meant that the Fe(III) heme affinities of these mutants are sufficient for them to acquire heme and to process pri-miRNAs. In contrast, all heme-binding-deficient mutants of DGCR8 we previously characterized are inactive in cells and have lower affinities for Fe(III) heme. We discovered that the Fe(III) heme affinity threshold for activating DGCR8 is shifted by overall heme availability changes in cells. These results indicate the presence of an available ferric heme pool that distinctly determines pri-miRNA processing efficiency in cells. Our study suggests cellular redox state and currently unknown Fe(III) heme-specific transporter proteins may be important for regulating miRNA maturation.

Secondary structure is a defining characteristic of pri-miRNA substrates. Substantial effort has been made to identify the important features of processing-competent pri-miRNA hairpins. Little attention is paid to structures outside the canonical hairpin. We identified a helix (f-helix) flanking the pri-miR-30a hairpin. Using an *in vitro* processing assay, we found that

disrupting the pairing in this structure caused processing defects that could be rescued by repairing the helix. Although f-helices are not found in all pri-miRNAs, our finding has important implications on how to improve DNA vector-based RNA interference technology. The second generation short-hairpin RNAs (shRNA^{mir}) are designed to mimic pri-miRNA and thus produce small RNAs through the natural miRNA maturation pathway. Their structure is most often based on pri-miR-30a. Indeed, we showed that the f-helix was functionally important for shRNA^{mir} processing and knockdown efficiency. This work indicates that structural elements surrounding the pri-miRNA hairpin may serve as regulatory elements for miRNA maturation.

miRNA expression is globally repressed in many tumors, and this dysregulation can drive tumorigenesis. High numbers of somatic mutations are a hallmark of many tumors, but distinguishing between disease-driving and spurious mutations remains a challenge. We analyzed seven missense DGCR8 mutations and found that four are severely defective. We determined that the E518K mutation identified in about 3% of Wilms Tumors, most likely disrupts the folding of DGCR8. The F448L mutation found in colon cancers rendered DGCR8 unable to bind Fe (III) heme and the Drosha protein. This result suggests that heme-binding to DGCR8 may be required for strong association with Drosha. Biochemical characterization of K289E and G336E show they have reduced affinities both pri-miRNA and heme. Altogether, these results suggest that a substantial fraction of tumor-derived DGCR8 mutations results in functional deficits and thereby are likely to make important contribution to a cancer-promoting phenotype.

Overall, my thesis investigates the fundamental mechanism of pri-miRNA processing, and its applications to biotechnology and diseases. My main scientific contribution is the identification and characterization of the unique RNA-binding heme domain in DGCR8. My work demonstrates the importance of heme in pri-miRNA processing, and defines an Fe(III) heme pool that regulates MC activity, thereby linking miRNA biogenesis with heme biology and cellular redox environments. My research extends our understanding of the functional/regulatory

elements in pri-miRNA hairpins to include regions flanking the conserved pri-miRNA hairpin. Finally, I identify probable driver mutations for tumor formation by connecting several cancer-associated somatic point-mutants of DGCR8 to functional defects in processing.

The dissertation of Jen Elana Quick-Cleveland is approved.

Tracy L. Johnson

Alison Renee Frand

Reid C. Johnson

David S. Eisenberg

Feng Guo, Committee Chair

University of California Los Angeles

2017

Dedicated to the memory of my grandfather,

Curtis Quick.

TABLE OF CONTENTS

LIST OF FIGURES BY CHAPTER

LIST OF TABLES

ACKNOWLEDGEMENTS

The chapters in this doctoral thesis is work completed alongside many collaborators that contributed valuable experiments, feedback and discussion.

I think it is important to acknowledge the key people who push you to become a better version of yourself. Dr. Feng Guo stretched himself to be a good mentor to me. He often went well out of his comfort zone to meet me where I was comfortable. He took pains to be very honest with me. When I needed rest, he encouraged me to take it. When I was scared, he pushed me to find my confidence. Beyond this, he is still one of the only scientist I have known that approaches research with a completely open mind and almost no ego. It is a truly beautiful thing that I continue to try and emulate.

To my lab-mate Grant Shoffner, I thank you for your clear thinking. Conversations with you have helped me to organize my own thoughts and words to be a more careful version of myself. Further, you have been such a source of joy to me, and I feel very lucky to have worked with you.

To Dr. Tracy Johnson and fellow graduate student Dr. Tremylla Johnson, Dr. Princess Gilbert, soon to be Dr. Janet Charray, Dr. Diana Rigeur, and Dr. Safiyyah Ziyad who supported me mentally when I needed it. Our meetings were so important to me and to my success at UCLA.

To Dr. Jose Rodriguez who wrote me a MATLAB program to analyze and quantify fluorescence intensity because I was unhappy with the one I had. The hours of frustration you saved me must be acknowledged.

And lastly, I would like to acknowledge the UCLA academic community. When I interviewed at UCLA, what struck me most was the quality of the faculty. Six years later, I am no less inspired by the professors at UCLA. I have approached many PIs for advice and mentorship, all have been open and willing to help me. This is exemplified in the Boyer Hall community, where equipment, reagents and help are shared freely. I have used every lab on the

second floor like it was my own. I have gotten so much technical advice, and been taught by so many. Dan Anderson and Carol Eng have given me expert experimental advice many times over the years. Biological Chemistry department staff went out of their way to help me on more than one occasion. The Boyer hall building staff, always supportive, talked to me about how their experiences of seeing students go through their ups and downs, always telling me to keep pushing. Truly, the whole community embraced me. This open and supportive atmosphere was crucial to any successes I had here.

VITA

2005-2008	A.A. biology Cabrillo College, Aptos, CA
2008-2011	B.S. pre-med biology Cal State Monterey Bay, Seaside CA
2009-2011	McNair Scholar Program, Cal State Monterey Bay
2011-2017	Eugene Cota-Robles Fellowship, University of California Los Angeles
2011	NSF Competitive Edge Fellowship, University of California Los Angeles
2013-2016	NIH Cellular and Molecular Biology Training Grant, University of California Los Angeles
2014	Audree Fowler Fellowship, University of California Los Angeles

Publications

First author:

Quick-Cleveland, J., Jacob J.P., Weitz, S.H., Shoffner, G., Guo, F. The DGCR8 RNA-binding Heme Domain Recognizes Primary-miRNAs by Clamping the Hairpin. Cell Reports 7 (6),2014

Shoffner, G., Quick-Cleveland, J., Solorio, K.M., Pickard, C., and Guo, F. Tumor-associated Point Mutations in DGCR8 Disrupts Primary MicroRNA Processing. *In preparation*

Contributing author:
Weitz, S.H., Quick-Cleveland, J, Jacob, J.P., Barr, I, Senturia, R, Weiss, S, Guo, F. An Fe^{3+} heme pool regulates the Microprocessor Complex. *In preparation*

Presentations

RNA Society Meeting 2015
Madison, WI
Poster: "*Missense Mutations from Primary Tumor Tissues cause defects in microRNA processing"*

UCLA Molecular Biology Institute Retreat 2015
Los Angeles, CA
Oral Presentation: *"The primary step of microRNA biogenesis"*

UCLA Biological Chemistry Departmental Retreat 2013
Malibu, CA
Oral Presentation: *"The primary step of microRNA biogenesis"*

UCSF summer internship - undergraduate research symposium
San Francisco, CA
Oral/Poster: Application of Microfludic Chips on Neuronal Development

CSU Program for Education and Research in Biotechnology Symposium (CSUPERB) 2010
Santa Clara, CA
Poster: *"Searching for Bioactive Compounds in Marine Fungi"*

Society for the Advancement of Chicanos and Native Americans in Science (SACNAS) 2009
Dallas, TX
Poster: *"TM7, the uncultivable filamentous bacteria linked to periodontitus"*

Society for the Advancement of Chicanos and Native Americans in Science (SACNAS) 2008
Salt Lake City, UT
Poster: *"Towards Cultivation and Isolation of TM7 "*

Mentorship

HHMI mentorship program – (Dec. 2015 – June 2016)

Kristina Solorio – Undergraduate (2014 -2016)

*Student awarded CARES Fellowship for summer 2015,

accepted to Princeton Molecular Biology Graduate Program 2016

Ruth Miranda -- Undergraduate (2014-2015)

*Graduated in 2015 – accepted to UCLA MD program

Timothy Huang – Amgen Scholar, (Summer 2016)

Ayesha Hamid – SMC transfer student, (Summer, 2017

Chapter 1: Overview

As scientific progress allows deeper mining of non-coding regions of the genome, one gains a broader appreciation of the complexity of living systems. Far from the ideas of the 1970s about non-coding regions as "junk DNA"(1), it has become clear that non-coding sequences can serve functional roles and are often central players in all biological processes (2–4). Splicing diversifies the human arsenal of proteins using regulatory small nuclear RNA (snRNA) (5–7). Some mRNA transcripts can have thousands of different splice forms with different biological activities (8,9). Long non-coding RNAs (lncRNA) also play crucial roles in developmental processes. For example, the lncRNA Xist coats one x-chromosome leading to its inactivation (10). Transfer RNAs (tRNAs) help assemble amino acids into polypeptides (11). Small Interfering RNAs (siRNA) play a crucial role in immunity (12,13). Small nucleolar RNAs (snoRNAs) are key players in processing and modification of rRNAs (14,15). PIWI-interacting RNAs (piRNAs) maintain genome stability in germ cells (2,16,17), and there are still more classes of non-coding RNA that we are learning about (miRtrons, circular RNA, small Cajal-body specific RNAs, competing endogenous RNAs), and surely more remain to be discovered (18).

MicroRNAs (miRNA) are a class of non-coding RNA that tune gene expression by negatively regulating mRNAs. Humans have ~2,600 miRNAs (19–22), which regulate an estimated 60% of all protein-coding genes (23). miRNAs are involved in nearly every aspect of development and cell physiology. miRNAs are single-stranded and ~22-nt in length. In the canonical biogenesis pathway miRNA are transcribed in the human body by RNA polymerase II, as long primary transcripts (pri-miRNAs). Maturation occurs through two cleavage (or processing) steps. The first processing step happens in the nucleus, where a conserved pri-miRNA hairpin is recognized by the Microprocessor Complex (MC). The core of the MC is the nuclease Drosha, and the RNA-binding hemoprotein DGCR8. Many other accessory proteins

can join the MC and modulate pri-miRNA processing. The cleavage produces precursor miRNAs (pre-miRNAs), which then exit the nucleus via the Exportin-5 protein (24–29). The second processing step is in the cytoplasm, where another ribonuclease Dicer cleaves pre-miRNAs to produce a miRNA duplex and helps unwind and load the mature miRNA strand onto the RNA-Induced Silencing Complex (RISC) (30–33). The loaded RISC uses the mature miRNA as a targeting molecule to bind specific mRNA, leading to translation repression or mRNA degradation (34,35). The mature miRNA binds to cognate sequences in the 3' untranslated region (3'UTR) of target mRNAs (24,27,36,37). Dysregulation in miRNA biogenesis leads to developmental defects (38–42), neurological disease(43–50), and cancer (51–58).

This work will focus on the nuclear step of miRNA biogenesis. The nuclear processing step defines the mature miRNA sequence by generating one end of what will become the mature miRNA. MC cleavage also determines the substrate for Dicer to cleave, generating the other end of the miRNA duplex. This is true for many miRNA, although a some individual miRNAs have been shown to have monouridylated and adenylated ends (59,60) While miRNAs are found in most organisms, the nuclear processing step mediated by the MC is specific to animals (25,61). Deletion of either Drosha or DGRC8, leads to embryonic lethality in mice, highlighting the essential role of miRNA in animal biology (41,62) . As such, it is not surprising that this pathway is under many layers of regulation.

Drosha is a RNase III family endonuclease that cuts both sides of the pri-miRNA hairpin with two RNase III domains to produce a characteristic two-nucleotide 3' overhang (26,30). Drosha localization is regulated by phosphorylation via GSK3β, where only phosphorylated Drosha is targeted to the nuclear compartment (63,64). The Drosha protein is stabilized by acetylation (the enzyme currently unknown), and by interaction with the protein TDP43, which increases processing activity (63,65).

The *DGCR8* gene resides in chromosome 22 in a region prone to deletion. One in four thousand live-births has a genetic defect called DiGeorge Syndrome or, more generally, 22q11.2 deletion syndrome (66). The size of the heterozygous deletion ranges from 1.5 million (8% of patients) – 3 million (80-90% of patients) base pairs and encompasses 30-60 genes (38,67). The phenotype of this deletion varies between individuals but some of the common symptoms include learning deficits, congenital heart disease, developmental defects in the palate, defects in the immune system and a 30% chance of developing schizophrenia later in life (38,67–69).

DGCR8 is at the crossroads of many regulatory pathways. The DGCR8 gene contains two pri-miRNA hairpins that are targeted by the MC leading to destabilization of DGCR8 mRNA, this creating a negative feedback loop (70,71). The DGCR8 protein is regulated post-translationally by phosphorylation, which stabilizes the protein (72). It has also been shown that DGCR8 can be deacetylated, which increases its affinity for pri-miRNAs (73). DGCR8 can be sumolyated and this enhances proteins stability as well as increases the MC's affinity for pri-miRNAs (74). Sequestration of DGCR8 by other proteins as well as other RNAs has also been observed (75). The pri-miRNA hairpin itself is marked by N^6-methyladenosine, which facilitates the recognition by DGCR8 (76). All these mechanisms intertwine to regulate the production of miRNAs.

My research focuses on four main questions concerning pri-miRNA processing.

Question 1 – how does DGCR8 recognize pri-miRNAs?

Pri-miRNAs contains a conserved hairpin secondary structure. Mainly this structure defines a pri-miRNA as a substrate for the MC. The canonical hairpin contains a ~30 bp stem with a terminal loop longer than 9 nucleotides (nt) (25,77,78). The base of the hairpin is bracketed by single-stranded regions (77). The transitions between single-stranded to paired

nucleotides, located at both ends of the stem, are termed pri-miRNA junctions. The protein domains that are responsible for recognizing these conserved features were not known. Several short sequence motifs are found in some pri-miRNA hairpins, and have been shown to enhance processing activity by promoting association of the MC and its accessory proteins (79).

We thought there had to be an interesting molecular mechanism that allowed the MC to specifically recognition of pri-miRNAs over all the other RNAs abundantly present in the nuclear compartment. Many molecular recognition problems are solved by signaling sequences that mark molecules and target them to interacting partners, however there is no conserved sequence motif that appears in all pri-miRNAs. The secondary structure of the conserved hairpin is believed to be the major marker driving pri-miRNA recognition. We wondered what specific role DGCR8 plays in recognizing pri-miRNAs.

Another unsolved issue is the function of heme in pri-miRNA biogenesis. In 2007, our group discovered that DGCR8 binds a heme cofactor (80). Further, we demonstrated that DGCR8 has a strong preference for Fe(III) heme, while Fe (II) heme-bound DGCR8 is inactive for pri-miRNA processing (81). Still other work in our lab demonstrated that DGCR8 ligates the heme iron using two cysteine side chains, the first time such a configuration ever been observed for a heme protein (82). DGCR8 binds heme as a dimer, with cysteine 352 from each subunit coordinating the iron in the porphyrin. Lastly, our data demonstrated that DGCR8 required heme for activity in cells (83,84).

My first major project in the lab is focused at elucidating the functional role of the heme-binding domain. DGCR8 contains two double-stranded RNA binding domains, and a C-terminal tail that is important for Drosha association. The N-terminus contains a nuclear localization sequence followed by the heme-binding domain (80,85). At the time, the “active DGCR8 core”

was defined as the two dsRBDs. There was a need to identify the function of the newly defined heme-binding domain and understand how it cooperates with other DGCR8 domains. I discovered that the heme-binding domain plays a critical role in recognizing the pri-miRNA junctions, also substantially increasing DGCR8's binding affinity for pri-miRNAs. Therefore, we renamed this region of DGCR8 the RNA-binding heme domain (Rhed) to better reflect its function. A year later, others confirmed the role of the Rhed in pri-miRNA binding and additionally found that the Rhed plays a key role in ensuring processing fidelity () and increasing the accuracy in pre-miRNA production. The Rhed represents the first known protein domain that binds both heme and RNA. Finally, this work implied that heme biology may have influence on pri-miRNA processing in human cells (84).

Question 2 – What is the functional role of DGCR8's heme cofactor?

Heme is an essential biological molecule for all living organisms. Heme serves as a cofactor for many proteins that drive important biochemical processes such as oxidative phosphorylation and oxygen transport (86–89). In normal physiology, a balance between synthesis, and degradation pathways tightly controls heme (90–92). Maintaining this balance is critical as high levels of free heme is toxic to cells due to its ability to produce reactive oxygen species through Fenton chemistry (93). Heme is synthesized in the mitochondria exclusively in the Fe(II) form, but only the Fe(III) form activates pri-miRNA processing by binding to DGCR8 (). It is unknown how readily available Fe(III) heme is in human cells. Based on characterization of the heme degradation enzymes, heme oxygenases (K_M = ~1 μM), it was proposed that the concentration of available heme may be just below the K_M. Other work in C. elegans suggests that available heme molecules are bound to heme transporter proteins. The authors propose there is almost no free heme in cells (94). Recent work using a ratiometric fluorescence heme sensor that primarily measures the Fe(II) form indicates that there is a pool of labile heme in *S. cerevisiae.* The concentration of heme is significantly higher (20-40 nM) in the cytosol than in either the

mitochondria or the nucleus (~2.5 nM) (95). Since DGCR8 activity is dependent on specifically Fe(III) heme we tackled questions about Fe(III) heme availability in the cell. We used two DGCR8 mutants; K446A, which is *active*, but has decreased heme affinity, and P351A, which is *inactive* and has even lower heme affinity. We found that, in HeLa cells cultured under a commonly-used heme-sufficient condition, the activity of DGCR8 mutants depends on whether their affinities for Fe(III) heme exceed a threshold. This threshold may be set by the intracellular available Fe(III) heme concentration. Alternatively, it could be imposed by a currently unknown heme transporter. Heme depletion from the media and inhibition of heme synthesis in cell culture reduce the threshold, further supporting our model. In combination with the inactive DGCR8 mutants that our group has characterized previously, our results suggest that Fe(III) heme availability in HeLa cells falls in a surprisingly narrow and low concentration range. These results demonstrate that mammalian cells maintain a well-defined and tightly-controlled pool of available Fe(III) heme important for miRNA maturation and possibly other fundamental physiological functions. This work draws connections between miRNA biogenesis, heme trafficking and redox regulation.

Question 3 – Do non-canonical structural features in pri-miRNAs affect processing efficiency?

Previous RNA secondary structure prediction, sequence conservation and mutagenesis analyses revealed the hairpin structural features of pri-miRNA. These advancements opened the possibility of designing pri-miRNA-mimics that could be programed to target any gene of interest. The design of DNA-vector-based RNA interference constructs, called short hairpin RNAs (shRNAs), has been closely following our understanding of the natural miRNA maturation pathway. The first-generation shRNAs were designed based on the hairpin structure of pre-miRNA (96). After the nuclear processing step was characterized (24–29), it was found that longer shRNA constructs that mimic endogenous pri-miRNA transcripts were better at knocking down genes (77,97,98). These second-generation shRNA vectors (called shRNA^{mir}) are

engineered based on natural pri-miR-30a sequences. Understanding the important elements that ensure efficient production of correct mature miRNAs also offer opportunities for improving the knockdown potency of shRNA^{mir}. Pri-miR-30a is among the few examples used in mechanistic studies in earlier years. It is processed efficiently and is by far the most widely used endogenous backbone for shRNA design. Interrogation of the pri-miR-30a hairpin for features that optimize mature miRNA production has been an aim of several research groups (99–101). Recently several pri-miRNAs have been compared to identify other optimal backbones (102). Several sequence motifs have been implicated in enhancing processing and at least one of these motifs appear in about half of all miRNA (79). These sequence motifs can compensate non-optimal hairpins. A potent single-copy shRNA vector called miR-30-E was recently designed that increased knockdown efficiency over the standard miR-30a backbone ~30 fold by including one of these motifs (100).

Our lab sought to define the minimal pri-miRNA hairpin by systematically truncating pri-miR-30a from the 5' and 3' sides. We discovered that ~150 nt were required for robust processing activity in vitro. Secondary structure predictions of the 150-nt pri-miR-30a suggested a peripheral helix formed by the 5'- and 3'-sequences flanking the canonical hairpin. Mutations in this flanking helix (f-Helix) had negative effects on pri-miRNA processing activity. We wondered whether this helix is important for pri-miR-30a processing. Biochemical processing assays indicated that disrupting the f-Helix from the 5' or 3' side caused aberrant processing *in vitro.* As pri-miR-30a remains the standard shRNA vector, we broadened our investigation by designing shRNA based on the 150-nt pri-miR-30a that targeted luciferase. My results indicate that disruption of this non-canonical helix does indeed affect processing activity and gene knockdown efficiency in human cells.

In the age of CRISPR technology, shRNAs remain an important tool for both research and potential therapeutics. They do not require any genetic manipulation to work. Groups of

genes can be targeted and the knockdown effect is not permanent. Potential “vaccines” for HIV and cancer are also an exciting extension of shRNA technology (103,104). A cancer vaccine called FANG currently in phase 3 clinical trials works by simultaneously expressing granulocyte-macrophage stimulating-colony factor (GMCSF) and an shRNA that targets the furin convertase (104). Knockdown of furin convertase leads to downregulation of two immunosupporessive growth factors (TFGß1, TGFß2) that have been correlated with tumor progression and poor patient outcomes. Understanding all the important variables in shRNA design will increase its utility as a tool for genetic manipulation or therapy.

Question 4 - Can point mutations in DGCR8 contribute to tumorigenesis?

The insight into the role of DGCR8 in pri-miRNA processing allowed us to ask more questions about DGCR8, pri-miRNA and disease. miRNAs are inextricably linked to cancer as they help regulate key processes such as differentiation, and proliferation. Global down-regulation of miRNAs had been observed in many forms of cancer (53,105–107). In 2007, this global repression was shown to enhance cellular transformation and tumorigenesis (105). Genomic alterations in ovarian, breast and melanoma cancer samples involve miRNA maturation genes (108), and 50% of miRNA genes are in cancer associated genomic regions (109). This suggests modification of miRNA expression is a part of tumor pathogenesis. miRNAs involved in cancer development can be classified as oncogenic (oncomiRs) if they target tumor suppressors (such as the miR-17~92 cluster) or as tumor suppressors if they regulate proteins involved in proliferation (such as miR-30a), but the role of specific miRNA can change in different cell types. In ovarian cancers, low levels of Drosha and Dicer mRNA correlated to lower expression and were predictive of poor patient outcomes (108). In breast cancer cells lines, expression of Drosha, DGCR8, and DICER along with other miRNA biogenesis accessory proteins are significantly repressed under hypoxia (54), which is another

well-known hallmark of tumors. In prostate cancer, DGCR8 has been linked to tumor progression by binding with the PTEN tumor suppressor. This activates Akt, thereby promoting proliferation (110). In the most common pediatric kidney cancer (Wilms tumors), mutations in DGCR8 and Drosha strongly altered miRNA expression, and whole-genome sequencing of 534 tumor samples showed that 15% had mutations in both Drosha and DGCR8 (111,112).

Given the large body of evidence linking dysregulation of miRNA biogenesis and tumorigenesis, we wondered if tumor-derived mutations in DGCR8 could be mechanistically linked to pri-miRNA processing defects. We selected six point mutations from the Catalogue of Somatic Mutations in Cancer (COSMIC) database (113). We also included a recurring DGCR8 point mutation from Wilms tumors (E518K). Four of these seven point mutations had clear negative impacts on pri-miRNA processing. We found this was due to changes in pri-miRNA and heme affinity. Importantly, we discovered that a colon cancer-derived DGCR8 mutant (F448L) is unable to bind heme, also loses its ability to bind Drosha. It has already been shown that shRNA-mediated knockdown of Dicer, Drosha, and DGCR8 leads to enhance cellular transformation and accelerated tumor formation. Therefore, we expect that point mutations that severely affect DGCR8 activity would also result in a “pro-cancer” phenotype. It has been proposed that cancer pathogenesis includes accumulation of a series of mutations that in combination alter normal cell physiology. Our study suggests that point mutations in DGCR8 could make a key contribution toward tumorigenesis.

Figures

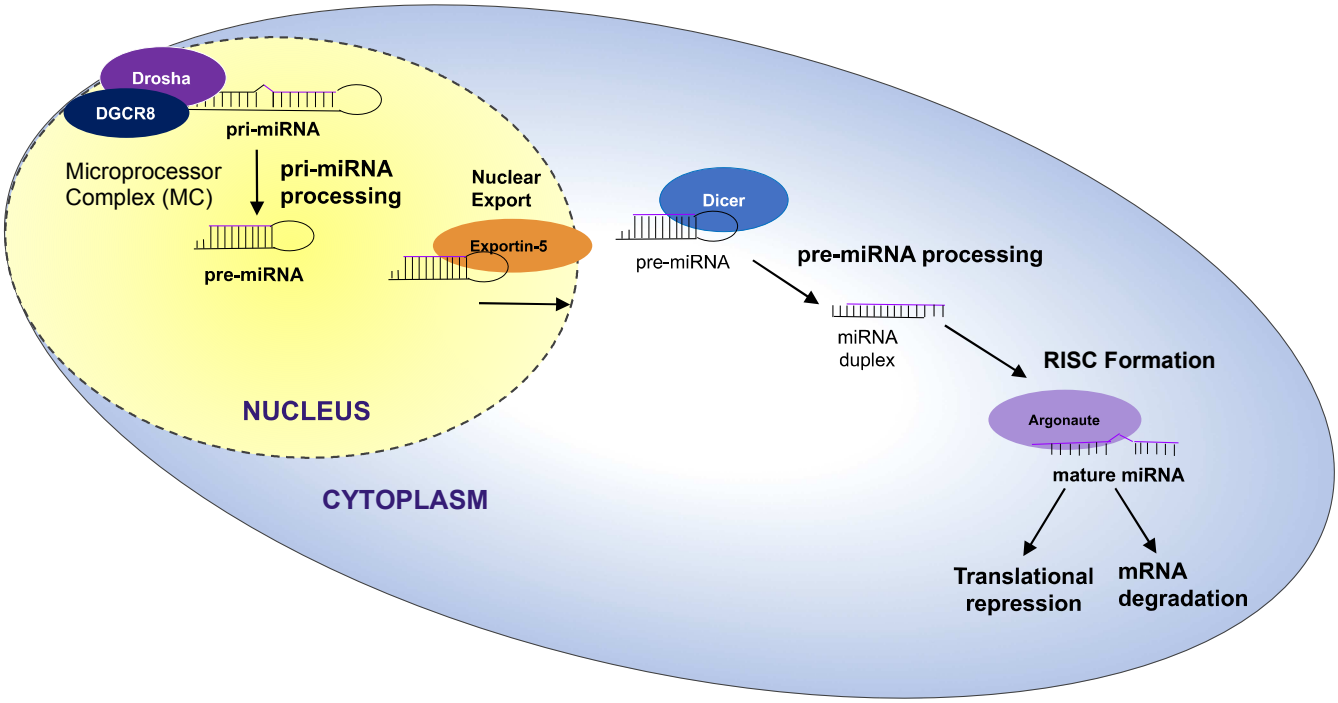


Figure 1. Canonical miRNA biogenesis pathway in animals. miRNAs are generated in a two-step maturation pathway. In the nucleus miRNA genes are transcribed as pri-miRNA. The MC cleaves the pri-miRNA to general pre-miRNA. This is exported from the nucleus where it is cleaved by Dicer in the cytoplasm. The miRNA duplex is loaded on Argonaute effector proteins forming RISC. Using the mature miRNA as a guide, RISC binds target mRNA leading to repression of the targets expression.

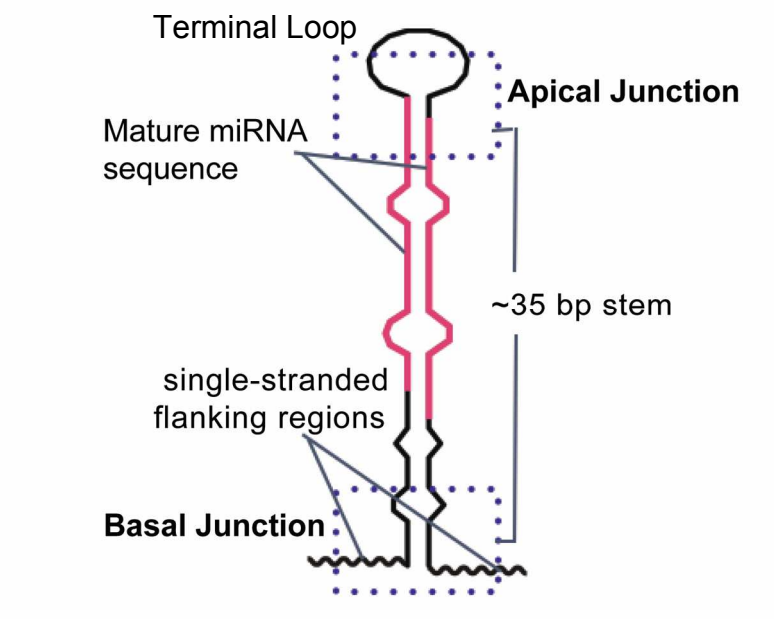


Figure 2. The structural features of a pri-miRNA hairpin.

References

1. Ohno S. So much “junk” DNA in our genome. Evol Genet Syst Brookhaven Symp Biol. 1972;23:366–70.

2. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet 2009;10(2):94–108.

3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature [Internet]. 2001;409(6822):860–921.

4. Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung W-K, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature [Internet]. 2007;447(7146):799–816.

5. Shuster, E.O., Guthrie C. Human U2 snRNA can function in pre-mRNA splicing in yeast. Lett to Nat. 1990;345:270–3.

6. Perriman, Rhonda, Ares Jr. M. Invariant U2 snRNA nucelotides for a stem loop to recognize the intron early in splicing. Mol Cell. 2010;38(3):416–26.

7. Merkhofer EC, Johnson TL. U1 snRNA rewrites the “script.” Cell [Internet]. Elsevier Inc.; 2012;150(1):9–11.

8. Sawaya MR, Wojtowicz WM, Andre I, Qian B, Wu W, Baker D, et al. A Double S Shape Provides the Structural Basis for the Extraordinary Binding Specificity of Dscam Isoforms. Cell. 2008;134(6):1007–18.

9. Einstein R, Jordan H, Zhou W, Brenner M, Moses EG, Liggett SB. Alternative splicing of the G protein-coupled receptor superfamily in human airway smooth muscle diversifies the complement of receptors. Proc Natl Acad Sci U S A [Internet]. 2008;105(13):5230–5.

10. Engreitz, J.M., Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X-chromosome. Science (80-). 2013;341(6147).

11. Loveland AB, Demo G, Grigorieff N, Korostelev AA. Ensemble cryo-EM elucidates the mechanism of translation fidelity. Nature [Internet]. Nature Publishing Group; 2017;546(7656):113–7.

12. Marques JT, Williams BRG. Activation of the mammalian immune system by siRNAs. Nat Biotechnol [Internet]. 2005;23(11):1399–405.

13. Robbins M, Judge A, Maclachlan I. siRNA and Innate Immunity. Oligonucleotides. 2009;19(2):89–101.

14. Decatur W a, Fournier MJ. RNA-guided nucleotide modification of ribosomal and other. RNAs, J Biol Chem. 2003;278(17):695–8.

15. Barth S, Shalem B, Hury A, Tkacz ID, Liang XH, Uliel S, et al. Elucidating the role of C/D snoRNA in rRNA processing and modification in Trypanosoma brucei. Eukaryot Cell. 2008;7(1):86–101.

16. Halic M, Moazed D. Transposon Silencing by piRNAs. Cell. 2009;138(6):1058–60.

17. Ma X, Wang S, Do T, Song X, Inaba M, Nishimoto Y, et al. Piwi is required in multiple cell types to control germline stem cell lineage development in the Drosophila ovary. PLoS One. 2014;9(3).

18. Cech TR, Steitz JA. The noncoding RNA revolution - Trashing old rules to forge new ones. Cell [Internet]. Elsevier Inc.; 2014;157(1):77–94.

19. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res.42:D68-73.

20. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 2011 Jan 1;39.

21. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res. Oxford University Press; 2008 Jan;36:D154-8.

22. Griffiths-Jones S. The microRNA Registry. Nucleic Acids Res. 2004 Jan 1;32(90001):109D–111.

23. Friedman RC, Farh KKH, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19(1):92–105.

24. Lee Y, Jeon K, Lee J, Kim S, Kim VN. MicroRNA maturation : stepwise processing and subcellular localization. 2002;21(17):4663–70.

25. Han J, Lee Y, Yeom K, Kim Y, Jin H, Kim VN. The Drosha – DGCR8 complex in primary microRNA processing. 2004;1–12.

26. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. The nuclear RNase III Drosha initiates microRNA processing. 2003;425(September):1–5.

27. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet [Internet]. 2004 Jul [cited 2013 Nov 7];5(7):522–31.

28. Gregory RI, Yan K, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature . 2004 Nov 11;432(7014):235–40.

29. Landthaler M, Yalcin A, Tuschl T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D . melanogaster Homolog Are Required for miRNA Biogenesis. 2004;14:2162–7.

30. Burger K, Gullerova M. Swiss army knives: non-canonical functions of nuclear Drosha and Dicer. Nat Rev Mol Cell Biol. Nature Publishing Group; 2015;16(7):417–30.

31. MacRae IJ, Zhou K, Doudna JA. Structural determinants of RNA recognition and cleavage by Dicer. Nat Struct Mol Biol. 2007 Oct 16;14(10):934–40.

32. Ma E, Zhou K, Kidwell MA, Doudna JA. Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. J Mol Biol. 2012 Sep 28;422(4):466–76.

33. Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X. A stepwise model for double-stranded RNA processing by ribonuclease III. Mol Microbiol. 2007 Nov 29 ;67(1):143–54.

34. Diederichs S, Haber DA. Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression. Cell. 2007 Dec 14;131(6):1097–108.

35. Schirle NT, Macrae IJ. The Crystal Structure of Human Argonaute 2. Science 336,1037(2012).

36. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004;116(2):281–97.

37. Triboulet R, Chang H-M, Lapierre RJ, Gregory RI. Post-transcriptional control of DGCR8 expression by the Microprocessor. RNA. 2009;15(6):1005–11

38. Kobrynski LJ, Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. Lancet 370, (9596):1443–52. 8

39. Pan Y, Balazs L, Tigyi G, Yue J. Conditional deletion of Dicer in vascular smooth muscle cells leads to the developmental delay and embryonic mortality. Biochem Biophys Res Commun. 2011 May 13;408(3):369–74.

40. Yi R, Pasolli HA, Landthaler M, Hafner M, Ojo T, Sheridan R, et al. DGCR8-dependent microRNA biogenesis is essential for skin development. Proc Natl Acad Sci U S A. 2009;106(2):498–502.

41. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet. 2007 Mar;39(3):380–5.

42. Kim H-J, Cho H, Alexander R, Patterson HC, Gu M, Lo KA, et al. MicroRNAs are required for the feature maintenance and differentiation of brown adipocytes. Diabetes. American Diabetes Association; 2014 Dec;63(12):4045–56.

43. Hsu R, Schofield CM, Dela Cruz CG, Jones-Davis DM, Blelloch R, Ullian EM. Loss of microRNAs in pyramidal neurons leads to specific changes in inhibitory synaptic transmission in the prefrontal cortex. Mol Cell Neurosci. 2012;50(3–4):283–92.

44. Raghavan V, Bhomia M, Torres I, Jain S, Wang KK. Hypothesis: Exosomal microRNAs as potential biomarkers for schizophrenia. Med Hypotheses.

45. Forstner AJ, Degenhardt F, Schratt G, Nöthen MM. MicroRNAs as the cause of schizophrenia in 22q11.2 deletion carriers, and possible implications for idiopathic disease: a mini-review. Front Mol Neurosci. 2013 Jan; 6:47.

46. Eom T-Y, Bayazitov IT, Anderson K, Yu J, Zakharenko SS. Schizophrenia-Related Microdeletion Impairs Emotional Memory through MicroRNA-Dependent Disruption of Thalamic Inputs to the Amygdala. Cell Rep; 2017;19(8):1532–44.

47. Monks S, Niarchou M, Davies AR, Walters JTR, Williams N, Owen MJ, et al. Further evidence for high rates of schizophrenia in 22q11.2 deletion syndrome. Schizophr Res. 2014 Mar;153(1–3):231–6.

48. Fénelon K, Mukai J, Xu B, Hsu P, Drew LJ, Karayiorgou M, et al. microdeletion, results in altered short-term plasticity in the prefrontal cortex. 2011;108 (11):4447-52.

49. Schofield CM, Hsu R, Barker AJ, Gertz CC, Blelloch R, Ullian EM. Monoallelic deletion of the microRNA biogenesis gene Dgcr8 produces deficits in the development of excitatory synaptic transmission in the prefrontal cortex. Neural Dev. 2011;6(1):11.

50. Stark KL, Xu B, Bagchi A, Lai W-S, Liu H, Hsu R, et al. Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. Nat Genet. 2008;40(6):751–60.

51. Iorio M V, Croce CM. Causes and consequences of microRNA dysregulation. Cancer J. NIH Public Access; 2012;18(3):215–22.

52. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. Nat Rev Cancer. NIH Public Access; 2015 Jun;15(6):321–33.

53. Volinia S, Calin GA, Liu C-G, Ambs S, Cimmino A, Petrocca F, et al. A microRNA

expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A. 2006;103(7):2257–61.

54. Bandara K V, Michael MZ, Gleadle JM. Hypoxia represses microRNA biogenesis proteins in breast cancer cells. BMC Cancer. 2014;14(1):533.

55. Hooda J, Mm A, Zhang L. Evaluating the Association of Heme and Heme Metabolites with Lung Cancer Bioenergetics and Progression. J Postgenomics Drug Biomark Dev. 2015;5(3).

56. Kota J, Chivukula RR, Donnell KAO, Wentzel EA, Montgomery CL, Hwang H, et al. Therapeutic microRNA Delivery Suppresses Tumorigenesis in a Murine Liver Cancer Model. Cell. Elsevier Inc.; 2009;137(6):1005–17.

57. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, et al. microRNAs exhibit high frequency genomic alterations in human cancer.

58. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nat Genet. 2007;39(5):673–7.

59. Heo I, Ha M, Lim J, Yoon M, Park J, Kwon SC, et al. Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. Cell. Elsevier Inc.; 2012;151(3):521–32.

60. Katoh T, Sakaguchi Y, Miyauchi K, Suzuki T, Kashiwabara S-I, Baba T, et al. Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev. Cold Spring Harbor Laboratory Press; 2009 Feb 15;23(4):433–8.

61. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. Nature Publishing Group; 2014;15(8):509–24.

62. Chong MMW, Zhang G, Cheloufi S, Neubert TA, Hannon GJ, Littman DR. Canonical and alternate functions of the microRNA biogenesis machinery. Genes Dev. Cold Spring Harbor Laboratory Press; 2010 Sep 1;24(17):1951–60.

63. Tang X, Li M, Tucker L, Ramratnam B. Glycogen Synthase Kinase 3 Beta Phosphorylates the RNAase III Enzyme Drosha at S300 and S302. 2011;6(6):1–6.

64. Fletcher CE, Godfrey JD, Shibakawa A, Bushell M, Bevan CL. A novel role for GSK3β as a modulator of Drosha microprocessor activity and MicroRNA biogenesis. Nucleic Acids Res.42(5).

65. Kawahara Y, Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. Proc Natl Acad Sci U S A. National Academy of

Sciences; 2012 Feb 28;109(9):3347–52.

66. Goodship J, Cross I, LiLing J, Wren C. A population study of chromosome 22q11 deletions in infancy. Arch Dis Child. BMJ Publishing Group;79(4):348–51.

67. Lindsay E a. Chromosomal microdeletions: dissecting del22q11 syndrome. Nat Rev Genet. 2001;2(11):858–68.

68. Maynard TM, Haskell GT, Peters AZ, Sikich L, Lieberman JA, LaMantia A-S. A comprehensive analysis of 22q11 gene expression in the developing and adult brain. Proc Natl Acad Sci U S A. National Academy of Sciences; 2003 Nov 25;100(24):14433–8.

69. Liu H, Abecasis GR, Heath SC, Knowles A, Demars S, Chen Y-J, et al. Genetic variation in the 22q11 locus and susceptibility to schizophrenia. 2002:99(26):16859-64.

70. Han J, Pedersen JS, Kwon SC, Belair CD, Kim Y, Yeom K, et al. Posttranscriptional Crossregulation between Drosha and DGCR8. Cell. Elsevier Inc.; 2009;136(1):75–84.

71. Shenoy A, Blelloch R. Genomic analysis suggests that mRNA destabilization by the microprocessor is specialized for the auto-regulation of Dgcr8. PLoS One,2009 Jan;4(9).

72. Herbert KM, Pimienta G, DeGregorio SJ, Alexandrov A, Steitz JA. Phosphorylation of DGCR8 Increases Its Intracellular Stability and Induces a Progrowth miRNA Profile. Cell Rep. 2013;5(4):1070–81.

73. Wada T, Kikuchi J, Furukawa Y. Histone deacetylase 1 enhances microRNA processing via deacetylation of DGCR8. Nat Publ Gr [Internet]. Nature Publishing Group; 2012;13(2):142–9

74. Zhu C,Chen C, Huang J, Zhang H, Zhao X, Deng R, et al. SUMOylation at K^{707} of DGCR8 controls direct function of primary microRNA. Nucleic Acids Res. 2015;43(16):7945–60.

75. Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, et al. Sequestration of DROSHA and DGCR8 by expanded CGG RNA Repeats Alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep. 2013;3(3):869–80.

76. Knuckles P, Carl SH, Musheev M, Niehrs C, Wenger A, Bühler M. RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat Struct Mol Biol.2017: 24(7):561-569.

77. Zeng Y, Cullen BR. Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences. 2005;280(30):27595–603.

78. Zeng Y, Yi R, Cullen BR. Recognition and cleavage of primary microRNA precursors by

the nuclear processing enzyme Drosha. 2005;24(1):138–48.

79. Auyeung VC, Ulitsky I, McGeary SE, Bartel DP. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell;152(4):844–58.

80. Faller M, Matsunaga M, Yin S, Loo J a, Guo F. Heme is involved in microRNA processing. Nat Struct Mol Biol [Internet]. 2007;14(1):23–9.

81. Barr I, Smith AT, Chen Y, Senturia R, Burstyn JN, Guo F. Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. Proc Natl Acad Sci U S A. 2012 Feb 7;109(6):1919–24.

82. Barr I, Smith AT, Senturia R, Chen Y, Scheidemantle BD, Burstyn JN, et al. DiGeorge critical region 8 (DGCR8) is a double-cysteine-ligated heme protein. J Biol Chem. American Society for Biochemistry and Molecular Biology; 286(19):16716–25.

83. Weitz SH, Gong M, Barr I, Weiss S, Guo F. Processing of microRNA primary transcripts requires heme in mammalian cells. Proc Natl Acad Sci U S A. 2014;111(5):1861–6.

84. Quick-Cleveland J, Jacob JP, Weitz SH, Shoffner G, Senturia R, Guo F. The DGCR8 RNA-Binding Heme Domain Recognizes Primary MicroRNAs by Clamping the Hairpin. Cell Rep. 2014 Jun 26;7(6):1994–2005.

85. Yeom K-H, Lee Y, Han J, Suh MR, Kim VN. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Res. 2006 Jan;34(16):4622–9.

86. Ponka P. Cell biology of heme. Am J Med Sci. 1999 Oct;318(4):241–56.

87. Tracz MJ, Alam J, Nath KA. Physiology and pathophysiology of heme: implications for kidney disease. J Am Soc Nephrol [Internet]. American Society of Nephrology;2007;18(2):414–20.

88. Mense SM, Zhang L. Heme: a versatile signaling molecule controlling the activities of diverse regulators ranging from transcription factors to MAP kinases. Cell Res. 2006;16(8):681–92.

89. Zhang, Li. (2011) Heme Biology. (Singapore: WORLD SCIENTIFIC).

90. Khan A a., Quigley JG. Control of intracellular heme levels: Heme transporters and heme oxygenases. Biochim Biophys Acta - Mol Cell Res: 2011;1813(5):668–82.

91. Ryter SW, Choi a M. Heme oxygenase-1: redox regulation of a stress protein in lung and cell culture models. Antioxid Redox Signal. 2005;7(1–2):80–91.

92. Ye W, Zhang L. Heme deficiency causes apoptosis but does not increase ROS generation in HeLa cells. Biochem Biophys Res Commun. 2004 Jul;319(4):1065–71.

93. Kohgo Y, Ikuta K, Ohtake T, Torimoto Y, Kato J. Body iron metabolism and pathophysiology of iron overload. Int J Hematol . Springer; 2008 Jul;88(1):7–15.

94. Severance S, Hamza I. Trafficking of heme and porphyrins in metazoa. Chem Rev. NIH Public Access; 2009 (10):4596–616.

95. Hanna DA, Harvey RM, Martinez-Guzman O, Yuan X, Chandrasekharan B, Raju G, et al. Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors. Proc Natl Acad Sci. 2016;113 (27):7539-7544.

96. Moore CB, Guthrie EH, Huang MT, Taxman DJ. RNA Therapeutics. 2010;629(2):1–15.

97. Bofill-De Ros X, Gu S. Guidelines for the optimal design of miRNA-based shRNAs. Methods. 2016;103:157–66.

98. Fowler DK, Williams C, Gerritsen AT, Washbourne P. Improved knockdown from artificial microRNAs in an enhanced miR-155 backbone: A designer's guide to potent multi-target RNAi. Nucleic Acids Res. 2015;44(5):1–16.

99. Myburgh R, Cherpin O, Schlaepfer E, Rehrauer H, Speck RF, Krause K-H, et al. Optimization of Critical Hairpin Features Allows miRNA-based Gene Knockdown Upon Single-copy Transduction. Mol Ther Nucleic Acids. 2014:3

100. Fellmann C ,Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, et al. An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. 2013;5(6):1704–13.

101. Wenwen F, Bartel DP. The menu of features that define primary microRNA and enable de novo desing of microRNA genes. Molecular Cell. 2015;131–45.

102. van den Berg FT, Rossi JJ, Arbuthnot P, Weinberg MS. Design of Effective Primary MicroRNA Mimics With Different Basal Stem Conformations. Mol Ther Acids. IOP Publishing; 2016;5(1).

103. Choi JG, Bharaj P, Abraham S, Ma H, Yi G, Ye C, et al. Multiplexing seven miRNA-Based shRNAs to suppress HIV replication. Mol Ther [Internet]. 2015;23(2):310–20.

104. Senzer N, Barve M, Kuhn J, Melnyk A, Beitsch P, Lazar M, et al. Phase I trial of "bi-shRNAi(furin)/GMCSF DNA/autologous tumor cell" vaccine FANG in advanced cancer. Mol Ther. Nature Publishing Group; 2012;20(3):679–86.

105. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing

enhances cellular transformation and tumorigenesis. 2007;39(5):673–7.

106. Ventura A, Jacks T. MicroRNAs and Cancer: Short RNAs Go a Long Way. Cell. 2009;136(4):586–91.

107. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. Nature [Internet]. 2005;435(7043):834–8.

108. Merritt WM, Lin YG, Han LY, Kamat AA, Spannuth WA, Schmandt R, et al. Dicer, Drosha, and Outcomes in Patients with Ovarian Cancer. 2008:359(25):2641-2650.

109. Zhang B, Pan X, Cobb GP, Anderson T a. microRNAs as oncogenes and tumor suppressors. Dev Biol. 2007;302(1):1–12.

110. Belair CD, Paikari A, Moltzahn F, Shenoy A, Yau C, Dall'Era M, et al. DGCR8 is essential for tumor progression following PTEN loss in the prostate. EMBO Rep [Internet]. 2015;16(9):1219–32.

111. Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, et al. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell.2015;27(2):286–97.

112. Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M, et al. Mutations in the SIX1/2 Pathway and the DROSHA/DGCR8 miRNA Microprocessor Complex Underlie High-Risk Blastemal Type Wilms Tumors. Cancer Cell. 2015;27(2):298–311.

113. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 2016;45:D777–83.

Chapter 2: The DGCR8 RNA-binding Heme Domain Recognizes pri-miRNAs by Clamping the Hairpin.

In this chapter, I present the results of a project investigating the heme-binding domain of DGCR8. Specifically, I was interested in the molecular mechanism that allowed the MC to recognize pri-miRNAs in the nucleus. This was my primary project. The initial observation that launched project was that the heme binding domain might participate in pri-miRNA recognition. I measured dissociation constants between the different domains of DGCR8 and pri-miRNA and discovered that the heme binding domain played a large role in binding pri-miRNA. Further, I found that the heme binding domain was the major specificity determinant allowing recognition of the proper RNA. This work presents the discovery of the RNA-binding heme domain (Rhed). I determined the stoichiometry of the DGCR8-pri-miRNA interaction, and discovered that the Rhed specifically recognized pri-miRNA junctions. Jose Jacob did the in vitro processing assays and designed some of the supplemental experiments. Sara Weitz performed the live-cell assays. Grant Shoffner had a parallel project with the goal of solving the structure of the DGCR8-pri-miRNA complex. He had been working with a zebrafish homologue of DGCR8 and he found that removing the base of the pri-miRNA hairpin was important for maintaining the binding stoichiometry of the complex. We published this work in Cell Reports.

Introduction

In the canonical microRNA (miRNA) maturation pathway in animal cells, miRNA primary transcripts (pri-miRNAs) are specifically recognized and cleaved by the Microprocessor to produce precursor miRNAs (pre-miRNAs) in the nucleus (Kim et al., 2009; Guo, 2012). Pre-miRNAs are exported to the cytoplasm, where they are cleaved by the ribonuclease Dicer, and mature miRNA strands are incorporated into the miRNA-induced silencing complexes. Previous studies have revealed fundamental features of pri-miRNAs, including a hairpin with mature miRNA strands located on either side of the stem (Lee et al., 2003) and unstructured regions

flanking the hairpin (Zeng and Cullen, 2005; Han et al., 2006). The hairpin stems contain roughly 30 bp with internal loops and bulges at variable positions. Base-pairing interactions in the stem, especially the bottom third, are clearly important for processing (Lee et al., 2003). The basal junction of a pri-miRNA, where the stem and the flanking unstructured regions join, is required for processing (Han et al., 2006). It has been proposed that the basal junction serves as an anchoring point for the Microprocessor to determine the cleavage sites ~11 bp away (the basal junction anchoring model). There have also been reports that apical region of the hairpin, including a ≥10 nt terminal loop, is important for processing (Zeng et al., 2005; Zhang and Zeng, 2010). Recently, three short (2–4 nt) sequence motifs have been shown to be enriched in nonnematode pri-miRNAs and important for processing of some pri-miRNAs in human cells (Auyeung et al., 2013). Overall, pri-miRNAs are defined primarily by their structures, with some sequence elements involved.

The Microprocessor has to identify true pri-miRNA substrates out of the myriad of other RNAs, and DGCR8 (*D*i*G*eorge *c*ritical *r*egion gene *8*, called Pasha in flies and worms) (Lee et al., 2003; Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004) plays a major role in this recognition. The 773-residue DGCR8 contains a nuclear localization signal (NLS) in the N-terminal region (Yeom et al., 2006; Shiohama et al., 2007), a central heme-binding domain, two double-stranded RNA-binding domains (dsRBDs), and a C-terminal tail (CTT) (Figure 1A). Prior to this study, the dsRBDs were shown to bind RNAs (Han et al., 2006; Faller et al., 2007; Sohn et al., 2007) and were defined as the “DGCR8 core” (Sohn et al., 2007). However, the dsRBDs alone bind pri-miRNAs with varying stoichiometry (Roth et al., 2013) and thus are unlikely to be the sole specificity determinant. DGCR8 has been shown to crosslink to pri-miRNA basal junction fragments (Han et al., 2006). However, the crosslinked protein and RNA residues have not been identified. It remains unknown how DGCR8 recognizes pri-miRNAs.

DGCR8 binds an essential heme (protoporphyrin IX in complex with iron) cofactor using a unique heme-binding domain. An active recombinant DGCR8 construct called NC1 (residues 276–751) and an isolated heme-binding domain (Figure 1A) both form constitutive dimers bound with one heme molecule (Faller et al., 2007; Senturia et al., 2010; Barr et al., 2011, 2012). Fe(III) heme directly binds the apo form of NC1 dimer and activates pri-miRNA processing in vitro (Barr et al., 2012). In HeLa cells, all known heme-binding-deficient DGCR8 mutants are inactive in pri-miRNA processing, and heme availability affects processing efficiency (Weitz et al., 2014). A WW-motif-containing dimerization subdomain (DSD) resides in the heme-binding domain and contributes a surface for heme binding (Senturia et al., 2010). Dimerization and heme binding appear to be conserved features of DGCR8 homologs (Senturia et al., 2012). Despite the body of evidence supporting the importance of the DGCR8 heme-binding domain, its function in pri-miRNA processing remains unclear.

In this work, we show that the heme-binding domain of DGCR8 plays a key role in pri-miRNA substrate recognition. Therefore, we have renamed this domain the RNA-binding heme domain (Rhed). The Rhed directly binds pri-miRNAs at the basal and apical junctions of the hairpin. Using cellular and biochemical pri-miRNA processing assays, we show that the Rhed and the Rhed-RNA interaction are important for DGCR8 activity. By collaborating with other domains, the Rhed allows full structural features of pri-miRNAs to be recognized.

Results

The Rhed of DGCR8 Directly Binds pri-miRNAs, Contributing to Affinity and Specificity

Our biochemical analyses indicate a function of Fe(III) heme-bound Rhed in pri-miRNA recognition. Because the Fe(II) heme-bound and heme-free forms of the human Rhed are insoluble at pH 5–8, we exclusively use the Fe(III) heme-bound Rhed dimer in this study and refer to this form of the protein as the Rhed. Filter-binding assays showed that the Rhed binds a panel of five pri-miRNAs (Figure S2) with K_d values ranging from 50 to 120 nM under equilibrium

conditions (Figure 1B; Table 1). These interactions appear to be specific to pri-miRNAs, as neither a 21 nt single-stranded RNA (ssRNA) nor an siRNA duplex can compete with pri-miR-21 for binding the Rhed in competition filter-binding assays (Figure 1C). tRNAs do compete, but not as efficiently as unlabeled pri-miR-21 (Figure 1C). Therefore, it is likely that the Rhed contributes to the pri-miRNA-binding specificity of DGCR8.

To estimate the relative contribution of the Rhed and dsRBDs to pri-miRNA binding, we measured the affinity of Rhed, NC1, and NC9 (=NC1 ΔRhed) for pri-miRNAs. The NC1 protein we used in this study is a Fe(III) heme-bound dimer unless stated otherwise. With deletion of the Rhed that is also responsible for dimerization, NC9 is a monomer. The affinity of the Rhed for each of the five pri-miRNAs (K_d = 50–120 nM) is comparable to that of the dsRBDs-containing NC9 (K_d = 30–70 nM) (Figure 1D; Table 1), indicating the importance of Rhed in RNA binding. Both the Rhed and NC9 have lower affinities for each of the pri-miRNAs compared to NC1 (K_d = 10–16 nM) (Figure 1D; Table 1), suggesting that these domains work together to achieve tight binding. The K_d values for NC1 and NC9 are in qualitative agreement with other measurements for similar constructs as previously reported (Faller et al., 2007, 2010; Sohn et al., 2007; Roth et al., 2013).

Each pri-miRNA Hairpin Contains Two Binding Sites for DGCR8

We analyzed the DGCR8-pri-miRNA-binding stoichiometry using size-exclusion chromatography (SEC), in which A_{450nm} (the Söret peak of the DGCR8-bound heme) and A_{260nm} (contributed mostly by RNA and to a lesser extent by the DGCR8-heme complex) were monitored simultaneously. Injection of NC1 mixed with pri-miR-23a or pri-miR-21 at 2:1 ratio resulted in single peaks (Figures 2A and 2B). Using a recently determined extinction coefficient (ϵ) of human DGCR8-bound heme (74 mM^{-1} cm^{-1} at 450 nm) (Senturia et al., 2012), we calculated the molar ratios of NC1 dimers and pri-miRNAs to be ~2:1 across the elution peaks (Figure 2B). These data suggest that there are two binding sites for DGCR8 dimers on each pri-

miRNA hairpin. The elution volume of the NC1-pri-miR-23a complex (8.3 ml) is close, but not identical, to the void volume (8.2 ml). We previously observed a similar elution volume for the NC1-pri-miR-30a complex (8.5 ml) (Figure 5D) (Faller et al., 2007). The 480 kDa apoferritin (one of the standard proteins used for calibration) also elutes in this region. These DGCR8-pri-miRNA complexes (including NC1-pri-miR-21) are expected to have molecular masses of ~260 kDa but greatly deviate from globular shapes and contain peripheral RNA strands of various lengths and structures. Therefore, it is not surprising that these complexes elute as if with higher molecular masses. Importantly, the protein-RNA ratios determined from the $A_{450}$and A_{260} measurements are independent of the elution volumes and shapes of the complexes.

The Rhed Determines the Stoichiometry of DGCR8-pri-miRNA Interaction

We performed similar SEC analyses using the Rhed. An input containing the Rhed and pri-miRNA (pri-miR-23a, pri-miR-21, or pri-miR-30a) at 2:1 molar ratio elutes in a single peak that is about 2 ml earlier than that of the free RNA (Figures 2C and S4A). Based on the A_{450} and A_{260} in the chromatograms, we calculated the Rhed:RNA ratio to be ~2:1 across the elution peak. Therefore, we conclude that there are two Rhed-binding sites on a pri-miRNA and that the Rhed is responsible for determining the DGCR8-pri-miRNA-binding stoichiometry. It is likely that the Rhed occupies similar pri-miRNA-binding sites whether it is in an isolated polypeptide or a part of processing-competent DGCR8 proteins.

We also analyzed the Rhed and pri-miR-30a complex at substoichiometric (1:1) input ratio. We observed an SEC elution peak at 11.4 ml, between those of the 2:1 complex (10.1 ml) and the free RNA (~12 ml) (Figure S4B). The Rhed:RNA ratio gradually changed from 2:1 to 0 across the elution peak, indicating the presence of multiple species at 2:1, 1:1, and 0:1 ratios that were partially resolved. This result suggests that, at least in the absence of the dsRBDs and CTT, the Rhed does not strongly prefer to bind one site versus the other.

The Rhed Binds Both Ends of a pri-miRNA Hairpin—The Apical and Basal Junctions

To locate the Rhed-binding sites on pri-miRNAs, we generated a series of truncated pri-miRNAs (Figures 2A and S3; Table S1) and analyzed their interactions with the Rhed using filter-binding assays and SEC. A pri-miR-23a truncation contains the 10 nt hairpin loop and 24 bp of the upper stem and thus includes the apical junction (aj-miR-23a-C; Figure S3A). The Rhed binds to aj-miR-23a-C with an affinity similar to that for pri-miR-23a (Table 1). Importantly, SEC analyses indicated that the Rhed dimer:aj-miR-23a-C molar ratio in their complex is reduced to ~1:1. When the Rhed:aj-miR-23a-C input ratio was 1:1, a single peak was observed, whereas increasing the input ratio to 2:1 resulted in a chromatogram containing both free-Rhed and complex peaks with nearly equal A_{450} (Figures 2D and 2E). We further truncated pri-miR-23a-C at its basal segments to generate aj-miR-23a-D, aj-miR-23a-E, and aj-miR-23a-F, which contain 20 bp, 11 bp, and 7 bp (Figures S3B–S3D), respectively. SEC analyses show that the Rhed forms ~1:1 complexes with all three RNAs (Figures 2F, 2G, S4C, and S4D), suggesting that the apical junction region contains a core binding site for the Rhed. Filter-binding assays showed that the affinity of Rhed for these RNAs decreases over the truncations, with K_d values increased to 191 nM for aj-miR-23a-D and the fractions of aj-miR-23a-E and aj-miR-23a-F bound to 1 μM Rhed only reached 0.25–0.30 and 0.15–0.20, respectively (Table 1). The large reduction in affinity caused by the deletion of 9 bp from aj-miR-23a-D suggests that the central segment of the pri-miR-23a stem is also important for the Rhed to bind the apical hairpin. Furthermore, to evaluate the importance of the hairpin loop, we replaced the 10 nt loop of aj-miR-23a-C with a GAAA tetraloop (aj-miR-23a-C-GAAA; Table S1) and found that the affinity of Rhed for this RNA decreased dramatically with the fraction of RNA bound to 1 μM Rhed only reached ~0.20. We also engineered pri-miR-21 to produce apical hairpins aj-miR-21-D and aj-miR-21-E that contain 18 bp and 10 bp in their stems (Figures S3E and S3F). The Rhed binds these RNAs with 1:1 stoichiometry (Figure S4E) and decreasing affinities over the truncations (Table 1), similar to the pri-miR-23a aj series. Therefore, we conclude that a Rhed-binding site

is located in the apical junction area and a high-affinity interaction requires both the hairpin loop and ~20 bp of the upper stem.

To identify the second Rhed-binding site, we deleted the apical hairpins from the pri-miRNAs and thereby produced basal junction models that include 8–9 bp of the lower stem and 7–9 nt of single-stranded regions on both sides of the hairpin (Figures 2A, S3G, and S3H). We linked the 5′ and 3′ strands using a GAAA tetraloop to stabilize the relatively short stem. The Rhed binds these bj-miRNAs, with affinities (K_d's of 232–326 nM) modestly lower than those for pri-miRNAs (K_d's of 50–120 nM) (Table 1), and the stoichiometry is ~1:1 (Figure 2H). As the bj RNAs are already quite short, we conclude that the second Rhed-binding site is located at the basal junction of a pri-miRNA.

Both the apical and basal Rhed-binding sites contain stem-ssRNA junctions, which are likely to be important features for Rhed to recognize. The stem and single-stranded regions of the junctions are all required for binding the Rhed, as neither ssRNA nor a duplex can compete with pri-miRNAs (Figure 1C) and substituting the hairpin loop of aj-miR-23a-C with GAAA disrupts the binding (see above). We also considered the possibility that certain pri-miRNA sequences drive the association with the Rhed. Three sequence motifs have recently been shown to be important for processing of some pri-miRNAs in human cells (Auyeung et al., 2013), including a "UG" at −14 position (14 nt upstream of the 5' Drosha cleavage site), a "UGU" or "GUG" at P22–P44 positions (22–24 nt into the pre-miRNA), and a "CNNC" at positions 16–17 (16–17 nt downstream of the 3' Drosha cleavage site). The −14 position is located at the basal junction, P22-P24 at the apical junction, and the 16–17 positions are close to the basal junction. We searched for these motifs in the five pri-miRNAs used in this study. pri-miR-30a has all three motifs, pri-miR-380 has none, and pri-miR-9-1, pri-miR-21, and pri-miR-23a each contain two (Figures S2). There is no clear correlation between the presence of these motifs and the affinity

for Rhed (Table 1). Therefore, it is likely that the Rhed recognizes the structures of pri-miRNA junctions.

The Rhed Is Required for pri-miRNA Processing

We next tested if the Rhed and its RNA-binding activity are important for pri-miRNA processing using cellular and biochemical assays. We first employed a recently developed fluorescent live-cell assay for pri-miRNA processing (Weitz et al., 2014). In this assay, a reporter plasmid inducibly expresses two fluorescent proteins, mCherry and eYFP (Figure 3A). A pri-miRNA sequence is inserted into the 3' UTR of the mCherry expression cassette, so that cleavage of the pri-miRNA reduces mCherry expression. The eYFP and mCherry fluorescent signals for individual cells have a linear relationship, and the slope faithfully indicates the efficiency of pri-miRNA processing but is not affected by subsequent steps of the miRNA maturation pathway (Weitz et al., 2014). Cotransfection of the reporter with the N-flag-DGCR8 expression plasmid (Figure 1A) increases pri-miRNA processing efficiency and the fluorescence slope (Figure 3B) and thus provides a robust method for measuring the activity of DGCR8 mutants. Endogenous DGCR8 is expressed at a very low level in HeLa cells and does not seem to interfere with the measurements as the N-flag-DGCR8 expression is typically 30- to 100-fold higher (Weitz et al., 2014). Using live-cell reporters containing either pri-miR-9-1 or pri-miR-30a, we found that deletion of the Rhed renders DGCR8 inactive. Unlike the wild-type, expression of N-flag-DGCR8 ΔRhed (Figure 1A) fails to increase the eYFP versus mCherry slopes relative to the transfections either without exogenous DGCR8 expression or with an inactive DGCR8 mutant, ΔCTT, in which the CTT is deleted (Han et al., 2004; Faller et al., 2010) (Figure 3B). The lost activity of ΔRhed is further supported by quantitative RT-PCR (qRT-PCR) measurements of the eYFP mRNA (for normalization), mCherry-pri-miRNA fusions, and mature miRNAs (Figures 3C and 3D) and is not caused by reduced DGCR8 protein expression or lack of nuclear localization (Figure 3E). Altogether, our data suggest that the Rhed is required for pri-miRNA processing in human cells.

Two previous studies showed that recombinant DGCR8 proteins without the Rhed and NLS are active for processing pri-miR-16 and pri-miR-30a in vitro (Yeom et al., 2006; Faller et al., 2007). To clarify the functional importance of the Rhed, we compared the processing activity of NC1 and NC9 (Figure 1A) in vitro using four additional pri-miRNA substrates (Figures S2A–S2D; Table S1). Deletion of the Rhed from the highly active NC1 abolishes processing of pri-miR-380, pri-miR-9-1, and pri-miR-21 at all DGCR8 concentrations tested in vitro (Figures 3F–3H and 3K). The activity of NC9 is greatly reduced for pri-miR-23a compared to that of NC1, with some processing observed at high NC9 concentrations (100 nM in Figure 3I and 200 nM in Figure 3K). We also confirmed that NC9 has substantial pri-miR-30a processing activity (Figures 3J and 3K), similar to the previous report (Faller et al., 2007). The pri-miR-30a sequence used in in vitro processing assays is identical to the insert of the pri-miR-30a cellular reporter. It is possible that, in vitro, recombinant DGCR8 constructs with the Rhed deleted *can* assemble with certain pri-miRNAs into productive complexes that are not formed in vivo. Altogether, our cellular and biochemical results demonstrate that the Rhed is important for DGCR8 function. These observations also highlight the importance of using both biochemical and cellular assays in studying pri-miRNA processing mechanism.

The RNA-Binding Surface of Rhed Is Important for pri-miRNA Processing

We next analyzed the RNA-binding surface of the Rhed using truncation and site-directed mutagenesis. The Rhed can be roughly divided into three regions: the N-terminal DSD, a central acidic loop, and a C-terminal region (Figure 1A). The DSD is soluble when expressed in *E. coli* without the rest of Rhed (Senturia et al., 2010). Filter-binding assays show that the DSD binds pri-miRNAs but does not bind the 21 nt ssRNA or the siRNA duplex. The affinities of the DSD for the five pri-miRNAs (K_d = 150–300 nM) are only modestly lower than those of the Rhed (K_d = 50–120 nM) (Table 1). However, unlike the Rhed, the DSD binds the junction-less aj-miR-23a-C-GAAA, with a K_d of 518 ± 45 nM (mean ± range, n = 2). These results suggest

that the DSD makes an important contribution to the Rhed-pri-miRNA interaction but does not retain all the affinity or exactly the same binding specificity.

Site-directed mutagenesis demonstrated that both the DSD and C-terminal regions of the Rhed contribute to pri-miRNA binding. The central loop is the least conserved among DGCR8 homologs and thus was not explored here. We previously determined crystal structures of the DSDs from human and frog DGCR8 (Senturia et al., 2010; Senturia et al., 2012). Inspection of the structures identified four surface-exposed basic residues (R322, R325, R341, and K342), which we mutated to alanine in pairs of spatial proximity (Figure 4A). The C-terminal region (residues 413–498) is rich in conserved basic residues (Senturia et al., 2012). Because no structure is available for this region, we systematically mutated them to alanine in groups with each containing two to three nearby mutations. These mutations were introduced to DGCR8 in a variety of contexts for pri-miRNA-binding and processing assays. Filter-binding assays indicated that the R341A/K342A mutations reduce the affinities by 3- to 5-fold for pri-miRNAs in the context of the NC1 and abolish pri-miRNA binding in the context of the Rhed (Table 1). The lack of pri-miRNA binding of Rhed R341A/K342A was confirmed using SEC analyses (Figure S5A). These results clearly indicate that R341 and K342 make a critical contribution to the Rhed-pri-miRNA interaction, most likely by directly participating at the RNA-binding interface. The modest affinity changes of the NC1 mutant are not surprising, as the dsRBDs are intact.

The R322A/R325A mutations reduce the affinity for pri-miRNAs by about 2- to 4-fold in the context of NC1 but make the Rhed protein insoluble and do not strongly alter the affinity for pri-miRNAs in the context of the DSD (Table 1). These observations suggest that R322 and R325 affect the Rhed-pri-miRNA interaction indirectly by partially disrupting the Rhed structure. The K424A/K426A/K431A mutations (G1) decreases the affinity for pri-miRNAs by 2- to 4-fold in the context of NC1 and render the Rhed protein incapable to plateauing to >50% occupancy of

most pri-miRNAs in filter-binding assays (Table 1). These RNA-binding defects were not caused by failure of the Rhed G1 mutant to bind the nitrocellulose membrane used in filter-binding assays, regardless whether pri-miRNAs are present (data not shown). However, SEC analyses of the Rhed G1-pri-miRNA complexes, assembled at higher protein and RNA concentrations (4 and 2 μM, respectively), showed peaks similar to those of the complexes formed by the wild-type Rhed protein (Figure S5B). Overall, our data suggest that at least some of the mutated residues contribute to pri-miRNA binding but that the RNA-binding defects of the mutant are not as severe as those of R341A/K342A.

The R341A/K342A, R322A/R325A and G1 mutants in the context of NC1 demonstrate reduced pri-miRNA processing activity in vitro (Figures 4B and 4C). Their electronic absorption spectra are indistinguishable from those of the wild-type (Figures S5C–S5F), ruling out the possibility that the pri-miRNA processing defects are caused by a heme-binding deficiency. Finally, these mutations were tested in the context of N-flag-DGCR8 using the live-cell reporter assay. The normalized eYFP versus mCherry slopes were 1.22 ± 0.03, 1.39 ± 0.04, and 1.44 ± 0.05 (±95% confidence interval, same below) for R322/R325A, R341A/K342A and G1, respectively (Figure 4D). These slopes are significantly lower (p values < 0.0001) than the 1.81 ± 0.06 for the wide-type N-flag-DGCR8 but also significantly higher (p values < 0.0001) than the 1.00 ± 0.02 for the reporter-only transfections. Immunoblotting analyses indicated that the N-flag-DGCR8 mutants were expressed at levels either similar to (R341A/R325A) or slightly lower than (R322A/K325A and G1) that of the wild-type (Figure 4E). We successfully compensated the lower expression levels of the N-flag-DGCR8 mutants by doubling the amounts of expression plasmids used in the transfections (Figure 4E) and observed no increase in the fluorescence slope for R322A/R325A (1.14 ± 0.05) and a slight increase of the fluorescence slope for G1 (1.57 ± 0.07) (Figure 4D). The G1 slope is still significantly lower than that of the wild-type control (1.88 ± 0.09) (p value < 0.0001). These results indicate that these mutations render the DGCR8 protein partially defective in cells. Altogether, our data demonstrate that the

DSD and the C-terminal region of the Rhed contribute to pri-miRNA binding and that the RNA-binding surface is important for the pri-miRNA processing activity of DGCR8.

Fe(III) Heme Is Required for Formation of Proper DGCR8-pri-miRNA Complexes

To understand the role of heme in DGCR8-pri-miRNA interaction, we analyzed the interaction between the heme-free apoNC1 proteins and pri-miRNA. We previously showed that apoNC1 has affinity for pri-miRNAs similar to that of the Fe(III) heme-bound form (Barr et al., 2012). This is not surprising, as the DSD and dsRBDs are still expected to be well folded. SEC analysis of the apoNC1 dimer with pri-miR-23a at an input ratio of 2:1 resulted in a peak at 10.8 ml (Figure 5A). This elution volume is between those of free pri-miR-23a (12.3 ml) and the Fe(III) heme-bound NC1-pri-miR-23a complex (8.3 ml), suggesting that the apoNC1 binds pri-miR-23a but in a conformation and/or stoichiometry different from those of the heme-bound NC1-pri-miR-23a complex.

We further analyzed pri-miRNA complexes with the apo form of a NC1 mutant P351A. The wild-type apoNC1 dimer is difficult to produce, usually has ~10% of residual heme associated, and is not very soluble at pH 7 and above (Barr et al., 2012). In contrast, NC1 P351A is easily purified as a heme-free dimer and is soluble at pH 8, at which the binding assays are performed. apoNC1 P351A can bind Fe(III) heme to reconstitute a complex similar to the wild-type (Barr et al., 2011). SEC of apoNC1 P351A dimer and pri-miR-23a at 2:1 input ratio resulted in a single peak at 10.7 ml, similar to the complex containing wild-type apoNC1 (Figure 5B). apoNC1 P351A with pri-miR-21 and pri-miR-30a also yielded elution peaks between those of the free RNAs and the heme-bound NC1-RNA complexes (Figures 5D and 5E). Increasing the protein:RNA input ratio to 3:1 did not shift the elution peaks, suggesting that the binding sites have been saturated (Figures 5C and 5F). Additionally, an increase of A_{280} was observed at ~14.2 ml, implying the presence of excess protein. Decreasing the input ratio to 1:1 shifted the elution peak to a larger volume, but not as far as that of the free pri-miRNA

(Figure 5G). These results suggest that apoNC1 binds a pri-miRNA with up to 2:1 stoichiometry. Overall, we conclude that the association of DGCR8 with Fe(III) heme causes a large conformational change in its complex with pri-miRNAs. It is also possible that the shift of elution volume is caused by partially disassembly of the apoNC1-pri-miRNA complexes during the SEC experiments. As heme-free DGCR8 is inactive in cells (Weitz et al., 2014), the conformation and/or stability induced by heme must be important for recognition and cleavage of pri-miRNAs by the Microprocessor.

The Covalent Linkage between the Rhed and dsRBDs Is Required for Proper Assembly of DGCR8-pri-miRNA Complexes

To further dissect the relationship between the Rhed and dsRBDs in pri-miRNA recognition, we tested if the two components of NC1, namely the Rhed and NC9, can mediate pri-miRNA processing in *trans* or assemble into proper tertiary complexes with pri-miRNAs. As already shown in Figure 3, with NC9 alone, pri-miR-9-1 cannot be processed and pri-miR-23a is weakly processed in vitro. NC9 and Rhed together do not alter the pri-miRNA processing activity comparing to the NC9 alone (Figures S6A and S6B), indicating that the covalent linkage between the Rhed and dsRBDs is important for pri-miRNA processing. In SEC analyses with both the Rhed and NC9, the pri-miRNAs eluted at volumes smaller than those of the Rhed-pri-miRNA binary complexes but different from those of the NC1-pri-miRNA complexes (Figures S6C–S6E), suggesting that some (nonproductive) tertiary complexes have formed. These observations suggest that the Rhed is responsible for properly anchoring the dsRBDs to pri-miRNAs for processing.

Discussion

Our study identifies the DGCR8 Rhed as the junction-binding domain that anchors the Drosha-DGCR8 complex to pri-miRNAs. It is expected that dsRBDs of DGCR8 associate with the stems of pri-miRNA hairpins (Sohn et al., 2007). The Rhed and dsRBDs together should

make extensive contacts with pri-miRNAs, allowing high-affinity binding and specific recognition. The Rhed-junction interaction provides a physical basis for previously proposed models regarding pri-miRNA recognition. For the basal junction anchoring model (Han et al., 2006), the Rhed binding to the basal junction allows the Microprocessor complex to measure ~11 bp and thereby determine the Drosha cleavage sites (Figure 6A). It was also suggested that the apical junction may serve as an alternative anchoring point for the Microprocessor and that this interaction leads to less efficient cleavages that are located in the middle of mature miRNA strands (called "abortive processing") (Figure 6A) (Han et al., 2006). The capability of the Rhed to bind the apical junction can explain the abortive processing. In a second model, the Microprocessor binds the hairpin loop (approximately equivalent to the apical junction) and measures ~22 bp to determine the Drosha cleavage sites (Zeng et al., 2005; Zhang and Zeng, 2010) (Figure 6B). For this "apical junction anchoring model," the Rhed-apical junction interaction results in productive processing. Common to both models is that association of DGCR8 with only one junction is sufficient to activate processing. The stoichiometry of DGCR8 and pri-miRNAs in processing competent complexes has been investigated in several studies, but no clear consensus has been reached. It has been previously reported that there is more than one copy of the DGCR8 protomer within the Microprocessor complex (Han et al., 2004). This observation may be explained by the dimerization of DGCR8 and/or by the formation of higher-order structures of DGCR8 upon binding pri-miRNAs. The crystal structure of the monomeric DGCR8 core showed that the two domains adopt an α-β-β-β-α fold typical for dsRBDs and that these domains are held together by a relatively rigid interface involving an extra α helix at the C terminus of dsRBD2 (Sohn et al., 2007). The RNA-binding surfaces of dsRBD1 and dsRBD2 are expected to point to distinct directions. Based on this and other evidence, it was suggested that each DGCR8 core binds either one pri-miRNA hairpin that is severely bent or two separate hairpins. A recent nuclear magnetic resonance and biochemical study indicated that the DGCR8 core binds pri-miR-16 fragments with varying stoichiometry and

that this interaction is reduced by competitor tRNAs (Roth et al., 2013). It was thereby concluded that the DGCR8-pri-miRNA interaction is nonspecific.

In contrast to the DGCR8 core, DGCR8 proteins containing the Rhed do bind pri-miRNAs with specificity (Han et al., 2006; Faller et al., 2010) (Figure 1C) and well-defined stoichiometry. Based on biochemical analyses of DGCR8-pri-miRNA complexes, we previously proposed that DGCR8 dimers form a cooperative higher-order oligomer upon binding to a pri-miRNA (Faller et al., 2010). Using a previously estimated ϵ_{450} of 58 mM^{-1} cm^{-1} for DGCR8-bound Fe(III) heme, we derived NC1 dimer:pri-miRNA ratios of ~3:1 in SEC peaks of their complexes (Faller et al., 2007, 2010). The ϵ_{450} was recently revised to 74 mM^{-1} cm^{-1}, using a widely accepted pyridine hemochromogen method (Senturia et al., 2012). Using this ϵ_{450} value, here we consistently obtained an NC1 dimer:pri-miRNA ratio of ~2:1 for the previous and new data. Our study suggests that the Rhed is a major specificity and stoichiometry determinant. Any mechanism regarding pri-miRNA recognition must include the dimeric Rhed of DGCR8. Technically, it is important to use recombinant DGCR8 proteins with high heme content in the investigation (Barr and Guo, 2014).

The observation of simultaneous and cooperative binding of two NC1 dimers to a pri-miRNA suggests a “molecular clamp” model. In this model, two DGCR8 dimers grip both ends of the hairpin using their Rheds and interact with each other, making extensive contacts with the pri-miRNA stem (Figure 6C). While it remains to be determined whether binding of both DGCR8 dimers to a pri-miRNA is required for assembly of active tertiary complexes with Drosha, this model is supported by recent cellular data showing that both the apical and basal junctions of pri-miRNAs are important for determining Drosha cleavage sites (Ma et al., 2013). Furthermore, a previous negative-stain electron tomography study yielded a “fat butterfly” density for a DGCR8-pri-miR-30a complex (Faller et al., 2010). Compared to the trimer-of-dimers proposal, the molecular clamp model seems to be more consistent with this electron density in that the

body of the butterfly may be the pri-miRNA hairpin and the four wings may be the DGCR8 subunits in the dimer of dimers.

The molecular clamp model allows the length of a full pri-miRNA stem to be measured through an interdimer interaction. The full stem length is one of the most important features of pri-miRNAs but is measured in neither the basal junction anchoring model nor the apical junction anchoring model. We previously showed that the CTT of DGCR8 contains an amphipathic α helix and that mutation of conserved hydrophobic resides on this α helix reduces binding cooperativity and abolishes pri-miRNA processing (Faller et al., 2010). This helix may be involved in the communication between the two DGCR8 dimers. The CTT is also required for binding Drosha (Han et al., 2006), possibly linking assembly of DGCR8 dimer of dimers to activation of RNA cleavage.

A DGCR8-pri-miRNA complex seems to contain built-in asymmetry, which should be able to help Drosha identify the correct cleavage sites that are closer to the basal junction. The apical junction of a pri-miRNA is next to a closed hairpin loop, whereas the single-stranded regions of the basal junction lead to the open ends of the RNA. This topological difference may contribute to the asymmetry of the complex. In fact, our data show different features in the Rhed interactions with the apical and basal junctions—the central region of the miRNA stem appears to be more important for Rhed association with the apical junctions than with the basal junctions (Table 1). Overall, the molecular clamp model unifies features of previous proposals, is consistent with most available experimental data, and explains how the full structural features of pri-miRNAs are recognized.

The mechanisms of substrate recognition by two ribonuclease III enzymes involved in miRNA maturation, Drosha and Dicer, bear interesting similarities and differences. Both Drosha and Dicer partner with dsRNA-binding proteins. However, Dicer-associated RNA-binding partners are not required for cleavage; instead, they modulate substrate affinity and cleavage

rates as well as loading of small RNAs to RNA-induced silencing complexes (Doyle et al., 2012). The PAZ domain of Dicer anchors the enzyme to the open end of a pre-miRNA hairpin, allowing the cleavage sites to be determined at a fixed distance (Macrae et al., 2006; Ma et al., 2012). The helicase domain of Dicer binds to the pre-miRNA hairpin loop and enhances the cleavage of pre-miRNAs over other Dicer substrates such as long dsRNAs (Tsutsumi et al., 2011; Ma et al., 2012). The interaction between the DGCR8 Rhed and pri-miRNA basal junctions seems functionally analogous to that between the Dicer PAZ domain and pre-miRNA open end. Similarly, the interaction between the Rhed and pri-miRNA apical junction may be comparable to the contact between the Dicer helicase domain and pre-miRNA hairpin loop. Thus, Drosha and Dicer systems appear to utilize distinct domains for the same purpose of recognizing the ends of substrate RNA helices.

The Rhed is an example of a heme-binding domain that directly binds nucleic acids. A number of transcription factors contain regulatory heme-binding domains, but these domains are separate from their DNA-binding domains (Gilles-Gonzalez and Gonzalez, 2005; Yin et al., 2007; Marvin et al., 2009). We believe that DGCR8 uses the heme cofactor for structural stabilization and/or regulatory functions. Without heme, DGCR8 still binds pri-miRNAs, but their complexes do not adopt processing-competent conformations. The Rhed appears to have evolved together with animal miRNAs, consistent with an essential function in pri-miRNA recognition. Neither the Rhed nor canonical miRNA is found in bacteria or archaea. Plants do not have the Rhed, and their miRNAs are processed from primary transcripts with longer hairpins by Dicer-like enzymes (Axtell et al., 2011). The Rhed is unique to DGCR8 homologs, whereas the dsRBDs are distributed among a wide range of organisms and in proteins involved in diverse biological functions (Masliah et al., 2013). Most canonical animal miRNAs are thought to originate from unstructured RNA sequences. Emergence of new canonical miRNAs requires successful processing and thereby the formation of junction-containing hairpin structures. Thus,

the Rhed of DGCR8 imposes a strong constraint for a new miRNA gene and serves as a gatekeeper for miRNA maturation and subsequent gene-regulation pathways.

DGCR8 has been shown to bind many other RNAs in mammalian cells, including mRNAs, small nucleolar RNAs, and long noncoding RNAs (Macias et al., 2012; Heras et al., 2013). In the inherited neurodegenerative disorder fragile X-associated tremor/ataxia syndrome, the expanded CGG repeats in the fragile X mental retardation 1 (*FMR1*) mRNA bind DGCR8, sequester the pri-miRNA processing machinery, decrease mature miRNA levels, and cause neuronal cell dysfunction (Sellier et al., 2013). Furthermore, DGCR8 and Drosha are required for the function of a class of artificial pri-miRNAs called shRNA^{mir}. As a DNA vector-based RNAi technology, shRNA^{mir} is widely used in biomedical research and is being explored for its therapeutic potential (Silva et al., 2005; Ni et al., 2011). Our characterization of the previously unknown RNA-binding domain in DGCR8 should aid understanding of its role in both miRNA and non-miRNA pathways and enhance the rational design of artificial pri-miRNAs in the future.

Figures

Table 1. Summary of K_d Values in nM Units Measured Using Filter-Binding Assays

pri-miRNA	Rhed	NC9	DSD	NC1
pri-miR-380	87 ± 7	33 ± 4	143 ± 13	12 ± 4
pri-miR-9-1	119 ± 18	52 ± 19	152 ± 51	10 ± 1
pri-miR-21	50 ± 7	45 ± 6	158 ± 10	12 ± 4
pri-miR-23a	121 ± 7	69 ± 0.6	208 ± 23	10 ± 1
pri-miR-30a	75 ± 6	72 ± 19	204 ± 3	16 ± 1

K_d of Rhed for Truncated pri-miRNAs

Apical Junctions	No. of bp in Stem	K_d	Basal Junctions	No. of bp in Stem	K_d
aj-miR-23a-C	24	102 ± 6	bj-miR-23a	9	232 ± 15
aj-miR-23a-D	20	191 ± 24			
aj-miR-23a-E	11	FB_{max} = 0.2–0.5			
aj-miR-23a-F	7	FB_{max} = 0.2–0.5			
aj-miR-21-D	18	178 ± 32	bj-miR-21	8	326 ± 32
aj-miR-21-E	10	FB_{max} = 0.2–0.5			

Mutations in the Rhed Region that Directly or Indirectly Affect the Affinity for pri-miRNAs

Mutants	R322A/R325A		R341A/K342A		K424A/K426A/K431A (G1)	
Context	NC1	DSD	NC1	Rhed	NC1	Rhed
pri-miR-380	25 ± 3	151 ± 16	32 ± 3	NB	50 ± 6*	57.3 ± 20
pri-miR-9-1	28 ± 5	116 ± 8	48 ± 5	NB	30 ± 5*	FB_{max} = 0.2–0.5
pri-miR-21	22 ± 1	240 ± 14	45 ± 3*	NB	19 ± 1*	FB_{max} = 0.2–0.5
pri-miR-23a	36 ± 20	128 ± 41	45 ± 5	NB	24 ± 5*	FB_{max} = 0.2–0.5

Rhed and dSRBDS Working Together To Achieve High-Affinity binding to pri-miRNAs. Errors are SD from three to nine repeats except those marked by " *," which indicate ranges from two replicates. In the binding reactions, the fraction of RNA bound (FB) to protein generally plateaued to >0.9. Those reactions in which the maximal FB only reaches 0.2–0.5 are marked as "FB $_{max}$ = 0.2–0.5." The highest protein concentrations used were 1 or 3 μM. See also Figures S2 Figures S2 and S3 S3and Table S1 Table S1 . NB, no binding.

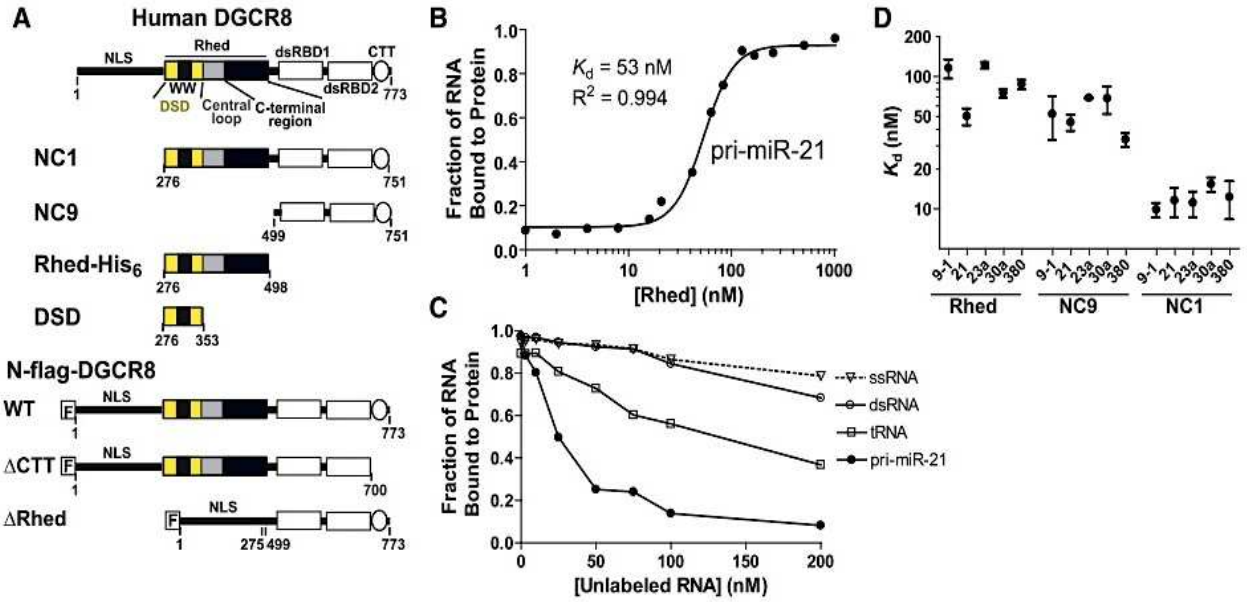


Figure 1. The Rhed Contributes to pri-miRNA Recognition by Directly Binding These RNAs and by Collaborating with the dsRBDs. (A) Recombinant human DGCR8 proteins used. "F" represents a FLAG tag. (B) A representative curve from filter-binding assays showing that the Rhed binds pri-miRNAs. The data were fit using a cooperative-binding model. The K_d is defined as the Rhed dimer concentration at which half-maximal RNA binding is achieved. (C) Competition filter-binding assays using unlabeled ssRNA, siRNA duplex, yeast tRNAs, or pri-miR-21 to compete with a trace amount of ^{32}P-labeled pri-miR-21 for association with 150 nM of Rhed dimer. An average molecular mass of 25 kDa was assumed in calculating molar concentrations of tRNAs.

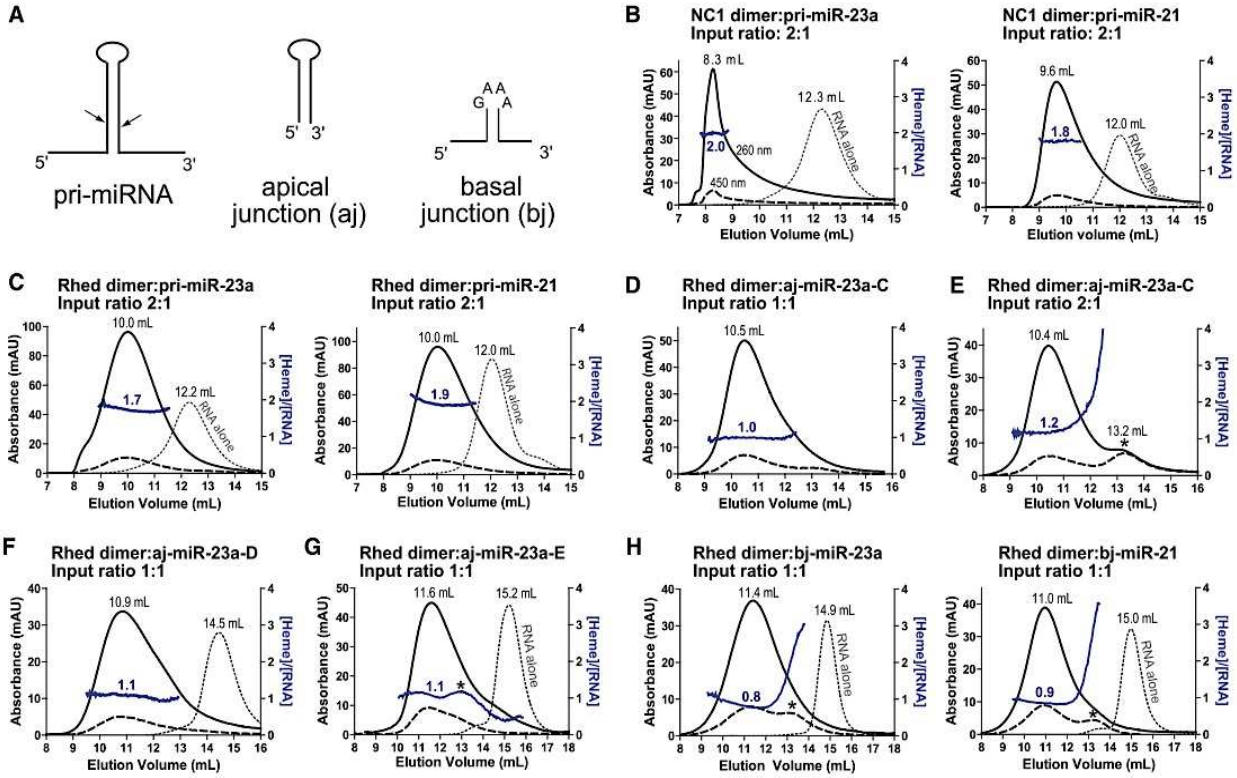


Figure 2. RNA Truncation and SEC Analyses Suggest that the Rhed Binds to pri-miRNA Junctions. (A) Schematics of pri-miRNA fragments. The arrows indicate the Drosha cleavage sites. The sequences and secondary structures are shown in Table S1 and Figures S2 and S3. (B–H) Size-exclusion chromatograms of NC1 in complex with 2 μM pri-miRNAs (B), the Rhed with 2 μM pri-miRNAs (C), the Rhed with 4 μM of aj-miR-23a-C at varying input ratios (D and E), the Rhed with 4 μM of aj-miR-23a-D (F), the Rhed with 4 μM of aj-miR-23a-E (G), and the Rhed with 4 μM of indicated basal junctions (H). Solid black lines indicate A_{260}, dashed lines show A_{450}, and dotted lines are A_{260} of the RNA-only runs. Solid blue lines represent heme:RNA ratios calculated from A_{450} and A_{260}, following the scale on the right y axis. The asterisk in (E), (G), and (H) marks a peak of free Rhed.

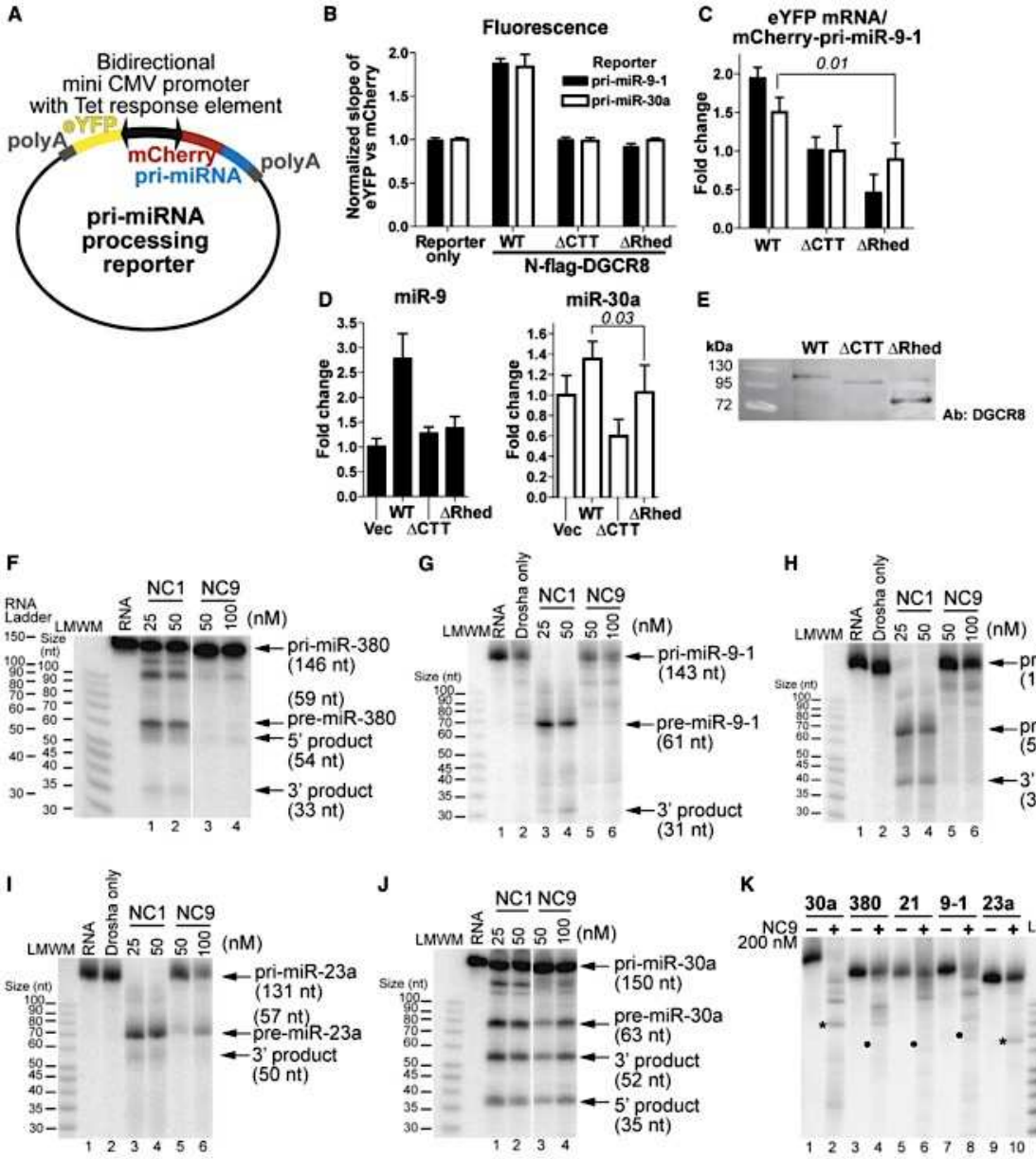


Figure 3. The DGCR8 Rhed Is Important for pri-miRNA Processing. (A) Schematic of the reporter plasmids. (B–E) The reporters were transfected into HeLa cells either alone or with the indicated N-flag-DGCR8 expression plasmids. (B) Slopes of the eYFP and mCherry fluorescence intensities, after normalization to that of the reporter-only transfection, are plotted. Error bars represent 95% confidence intervals. (C) Ratios of eYFP mRNA and mCherry-pri-miRNA (mean ± SD, n = 3 or 4). (D) Abundance of mature miR-9 and miR-30a normalized by that of β-actin (mean ± SD, n = 3). Select p values are indicated in italics. miR-30a is highly expressed endogenously in HeLa cells, and thus the relative changes are modest. (E) An anti-DGCR8 immunoblot of nuclear extracts from the transfected cells. An

equal amount of total proteins was loaded in each lane, as estimated using a Coomassie-stained SDS gel. (F–K) Reconstituted pri-miRNA processing assays. LMWM, low-molecular-weight marker. Relationship between LMWM and a true RNA ladder in 15% gels is shown in (F). In (K), the asterisks mark a pre-miRNA band and the dots mark the position expected for a pre-miRNA product

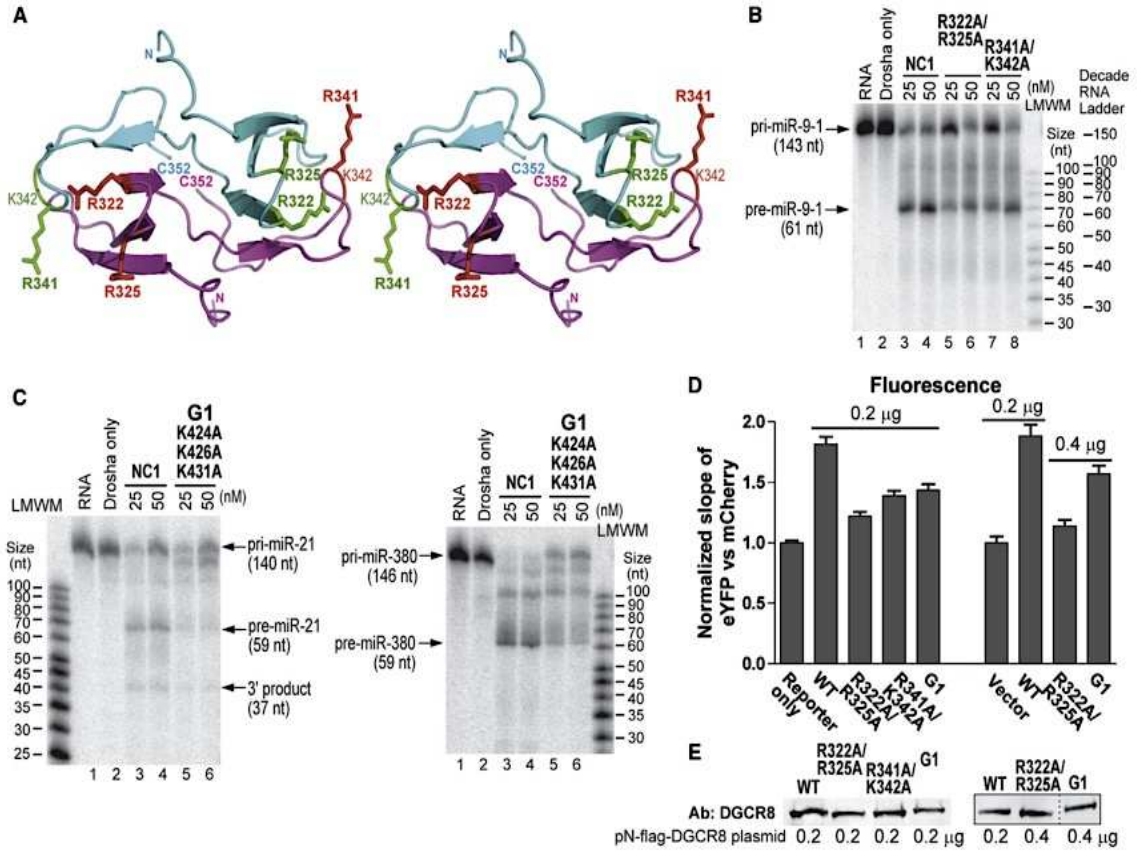


Figure 4. The pri-miRNA-Binding Surfaces of the Rhed Are Important for Processing.

(A) Stereo diagram of the DSD crystal structure of human DGCR8 (Protein Data Bank accession code 3LE4) (Senturia et al., 2010), with the side chains of the mutated residues shown in sticks. The two subunits are drawn in cyan and magenta. (B and C) Reconstituted pri-miRNA processing assays.(D) Cellular assays using the pri-miR-9-1 reporter. The amounts of DGCR8 expression plasmids or the pCMV-Tag-2A vector are indicated on the graph. Error bars represent 95% confidence intervals. The

presence of pCMV-Tag2A vector in the control transfection does not alter the fluorescence slope. (E) Anti-DGCR8 immunoblots of nuclear extracts from transfected cells. Equal amount of total proteins was loaded in each lane.

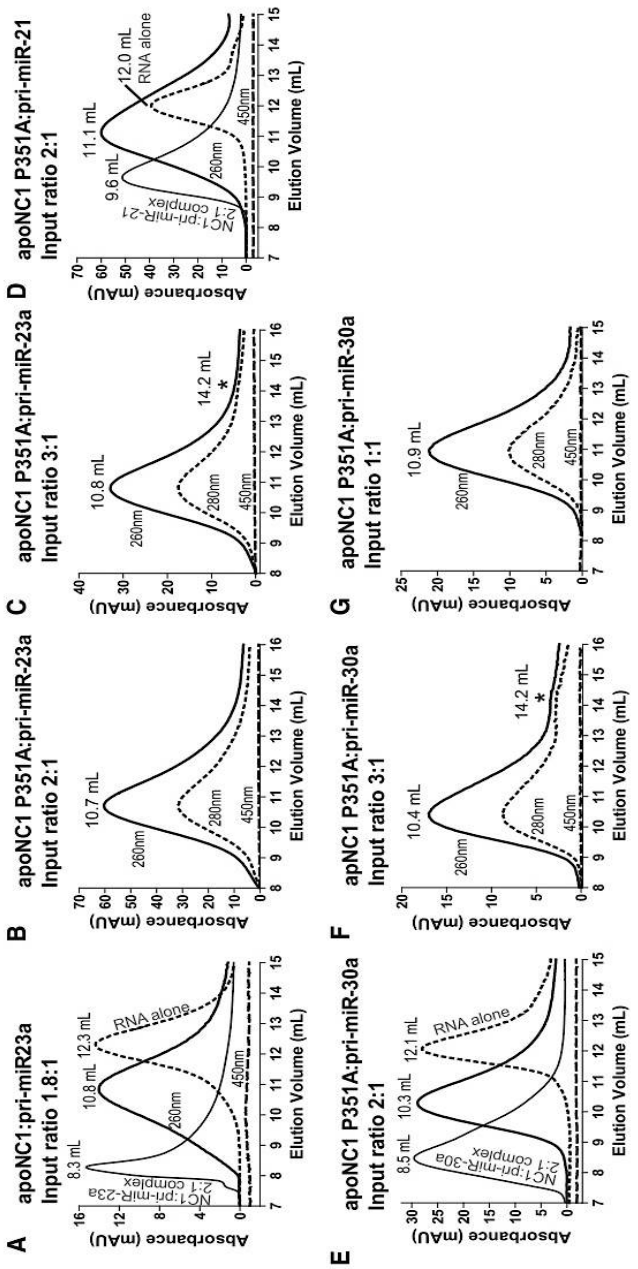


Figure 5. Fe(III) Heme Causes a Large Conformational Change to DGCR8-pri-miRNA Complexes. Size-exclusion chromatograms **of** (A) apoNC1 in complex with 0.45 μM pri-miR-23a and (B–G) apoNC1 P351A with 2 μM pri-miRNAs at the indicated input ratios. The asterisk in (C) and (F) marks a potential free protein peak.

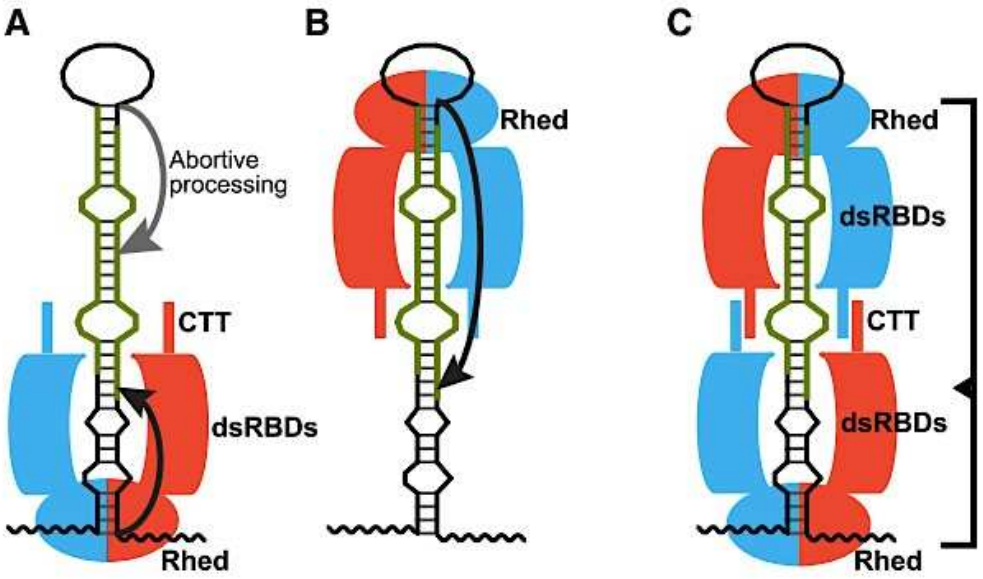


Figure 6. Models of How a pri-miRNA Is Recognized by the Microprocessor. (A) The basal junction anchoring model (Han et al., 2006). (B) The apical junction anchoring model (Zeng et al., 2005).(C) Our proposed molecular clamp model. See Discussion for details. The DGCR8 subunits in a dimer are shown in red and cyan. The thick avocado strands represent 5′ and 3′ mature miRNAs.

SUPPLEMENTAL FIGURES

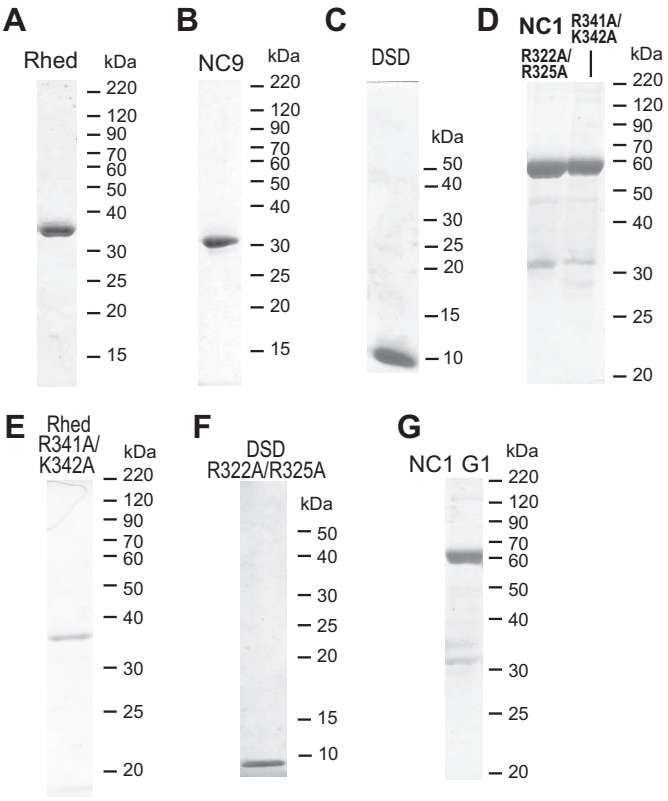


Figure S1. Coomassie-stained SDS-PAGE of purified recombinant DGCR8 proteins.

Related to Figures 1-4

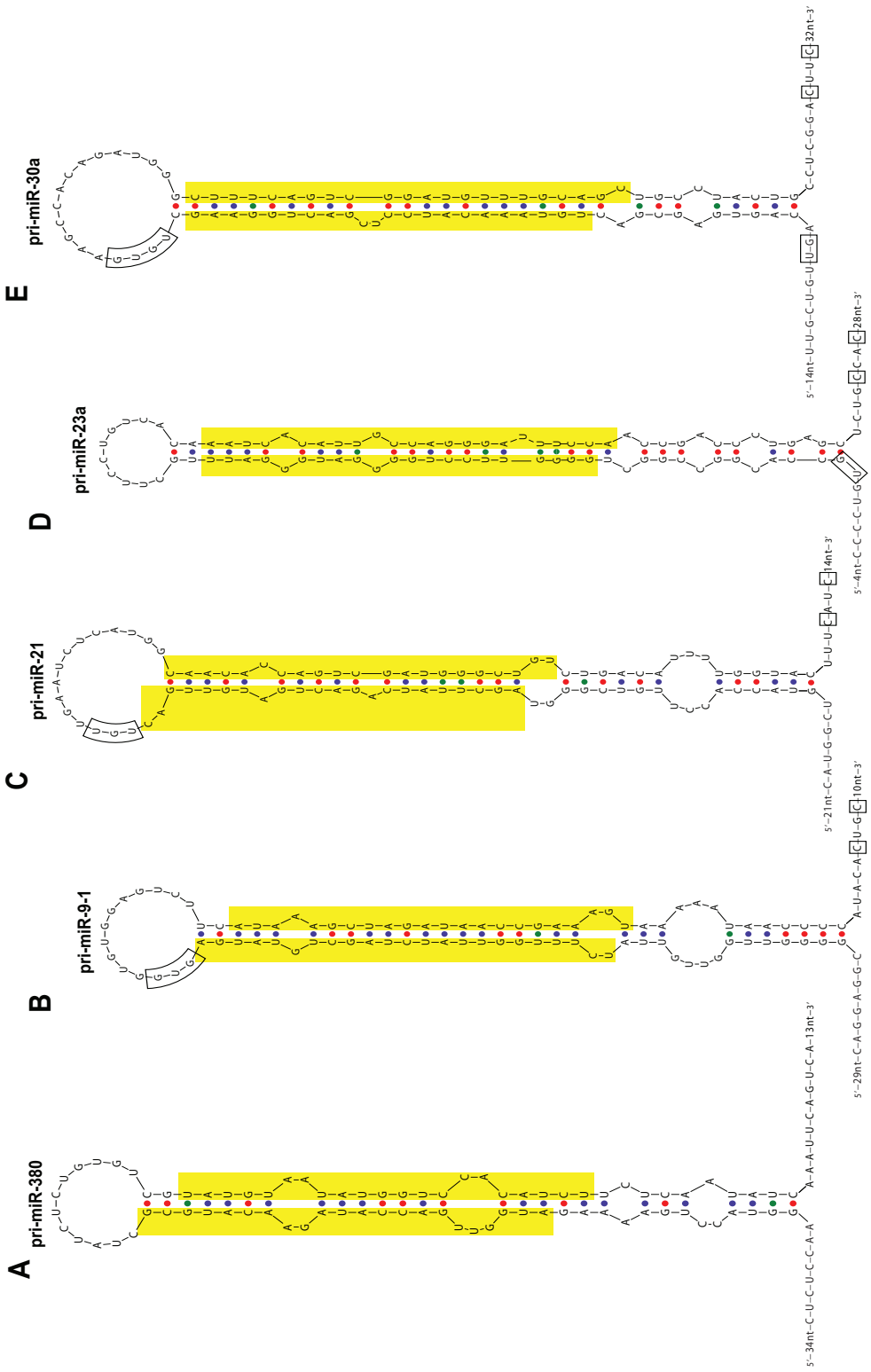


Figure S2. Secondary structures of pri-miRNAs. Related to Figures 1 and 2, and Table 1

(**A**) pri-miR-380. (**B**) pri-miR-9-1. (**C**) pri-miR-21. (**D**) pri-miR-23a. (**E**) pri-miR-30a.

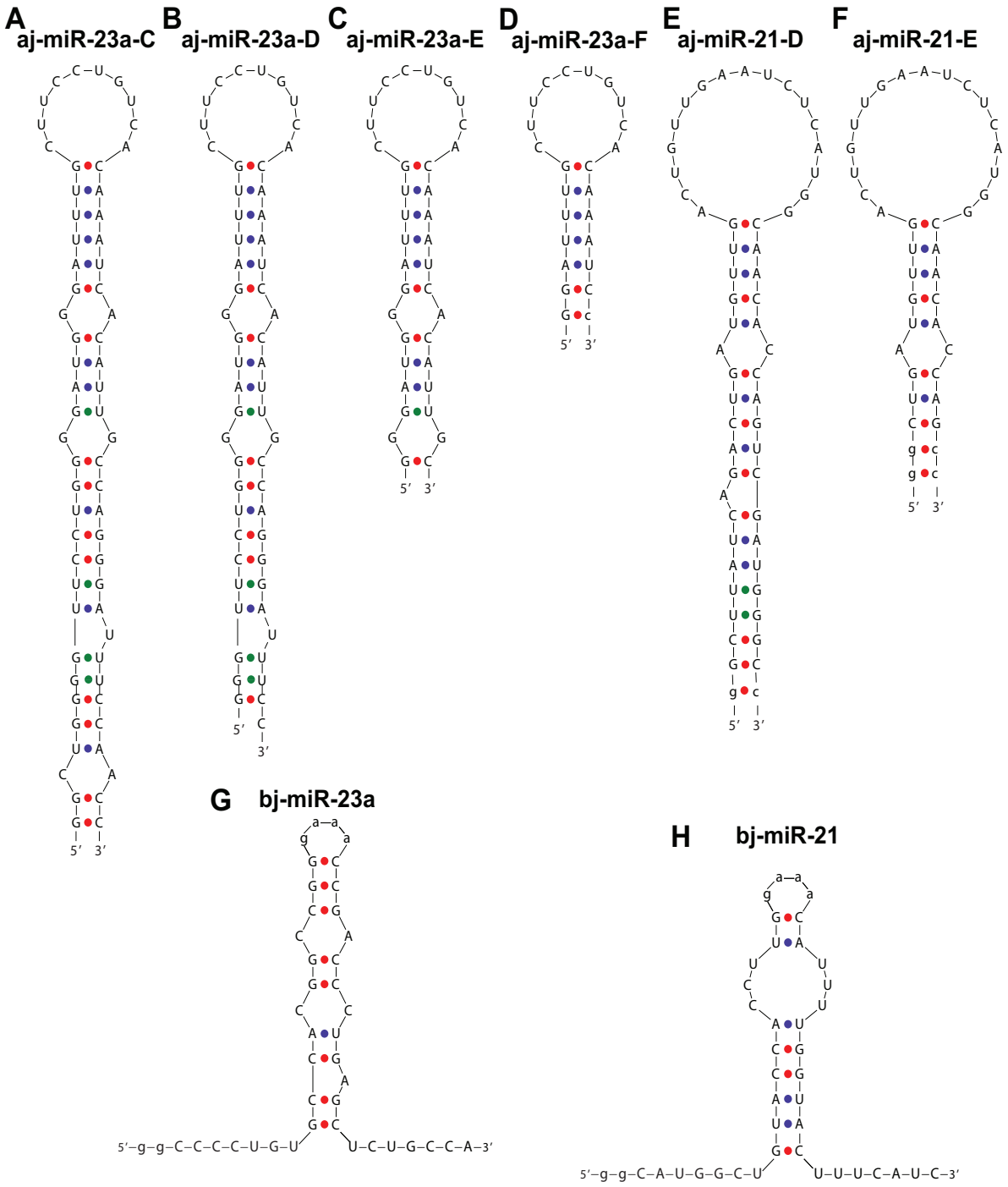


Figure S3. Secondary structures of pri-miRNA fragments. Related to Figures 1 and 2, and Table 1

(**A**) aj-miR-23a-C. (**B**) aj-miR-23a-D. (**C**) aj-miR-23a-E. (**D**) aj-miR-23a-F. (**E**) aj-miR-21-D. (**F**) aj-miR-21-E. (**G**) bj-miR-23a. (**H**) bj-miR-21. The structures were generated and graphed using MFOLD with the hairpin loop constrained to be single-stranded. Non-native residues that were introduced

to stabilize the structures or facilitate *in vitro* transcriptions are represented by lowercase letters. The motifs that were recently shown to be important for human pri-miRNA processing (Auyeung *et al*, 2013) are boxed in (**A**)-(**E**).

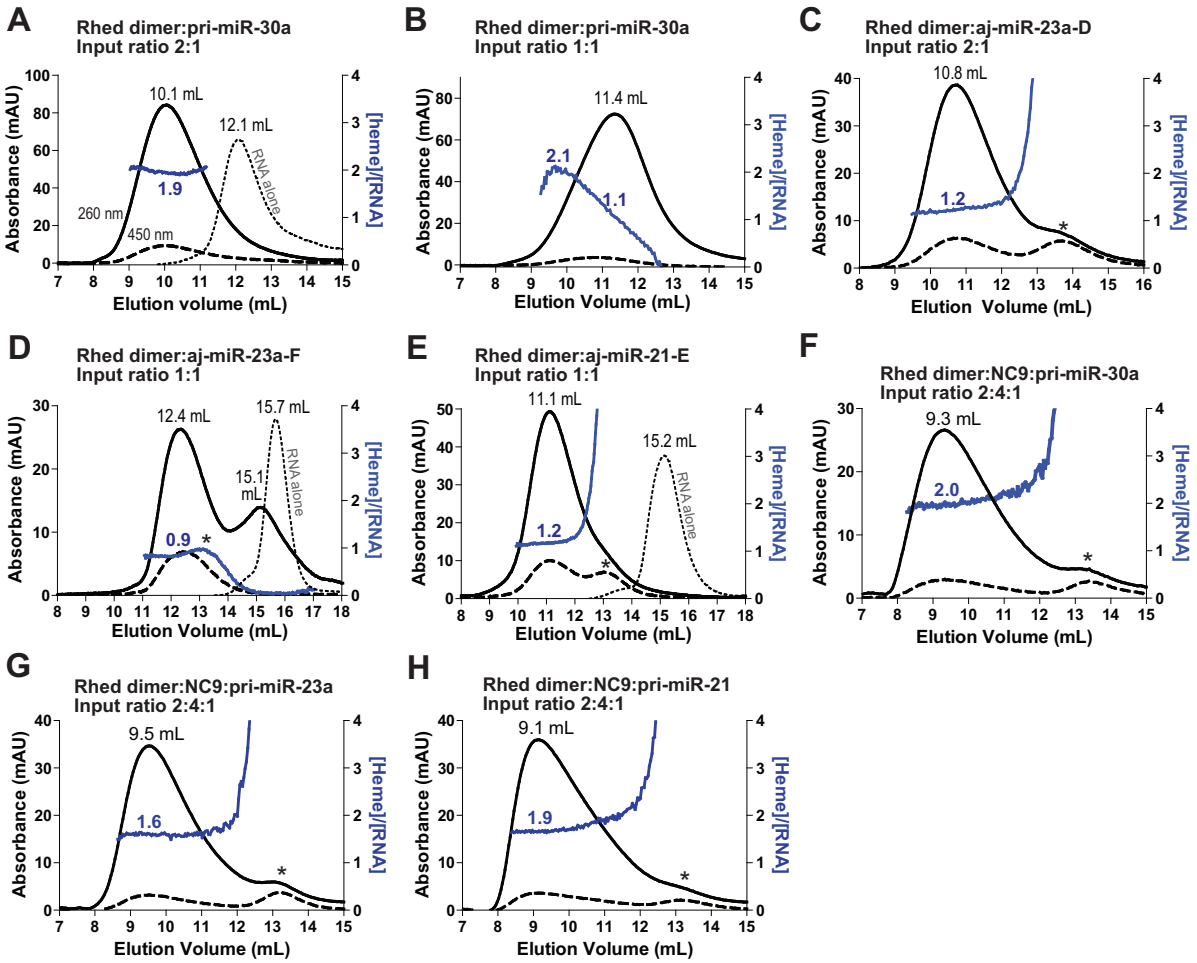


Figure S4. Size exclusion chromatograms of the Rhed in complex with pri-miRNAs. Related to Figure 2

The procedure and condition are similar to those used in Figure 2. The inputs contained 2 μM of pri-miR-30a (**A**,**B**), 4 μM of aj-miR-23a-D (**C**), 2 μM of aj-miR-23a-F (**D**), or 4 μM of aj-miR-21-E (**E**). Solid black lines indicate A_{260}, dashed lines show A_{450} and dotted lines are A_{260} of the RNA-only injections. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis. The asterisk indicates a free-RHED peak. The chromatogram of the RHED and aj-miR-23a-F displayed a minor A_{260} peak at 15.1 mL (corresponding to free RNA) and an increase of the heme:RNA ratio at 13.2 mL (hinting the presence of free RHED), indicating that the RHED has a low affinity for aj-miR-23a-F.

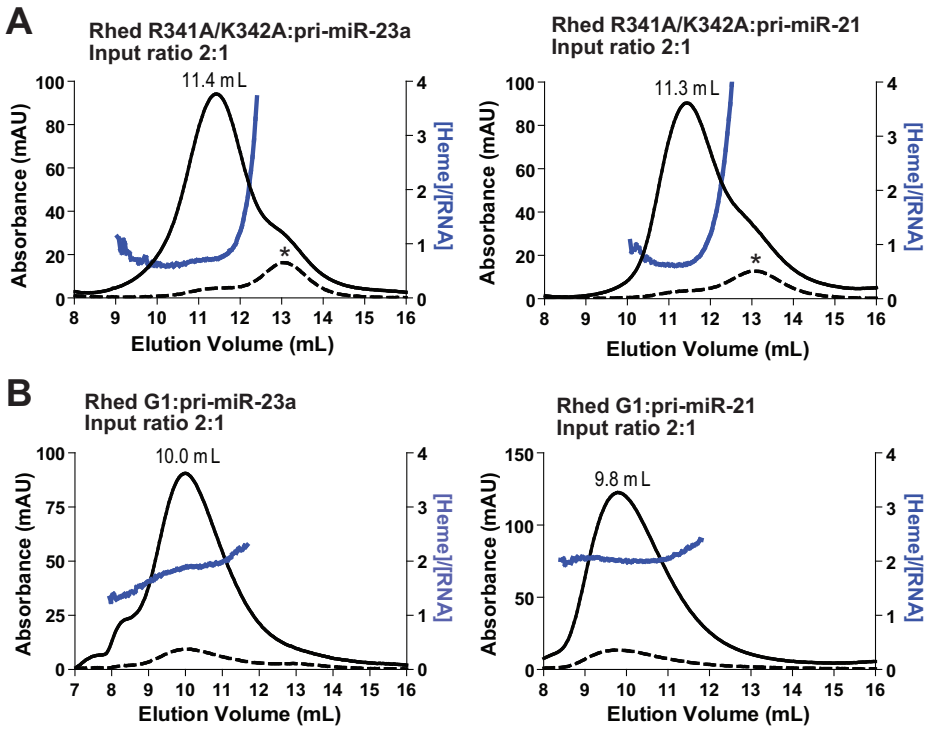


Figure S5. Size exclusion chromatograms of the Rhed R341A/K342A and G1 mutants in complex with pri-miRNAs. Related to Figure 4

The procedure and condition are similar to those used in Figure 2. The inputs contained 4 μM RHED dimer and 2 μM RNA. The asterisk indicates a free-RHED peak. Solid black lines indicate A_{260}, dashed lines show A_{450} and dotted lines are A_{260} of the RNA-only injections. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis.

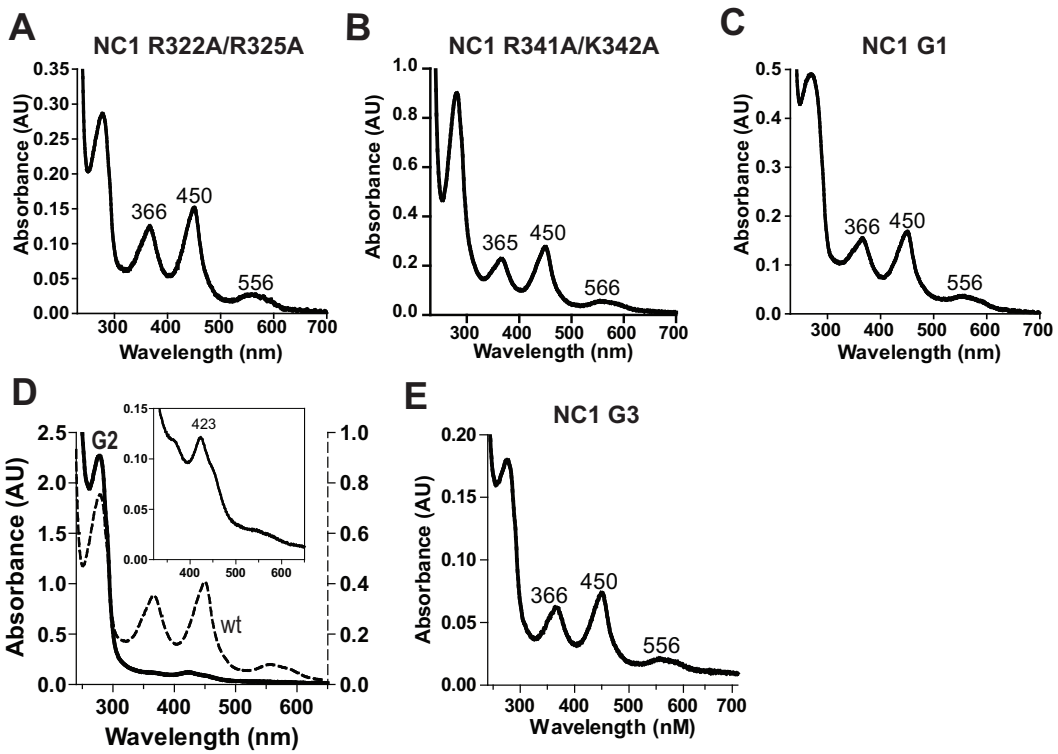


Figure S6. The NC1 R322A/R325A, R341 and G1 mutants bind heme similarly to the wild type. Related to Figure 4

Electronic absorption spectra of R322A/R325A (**A**), R341A/K342A (**B**), G1 (**C**) and wild type (**D**).

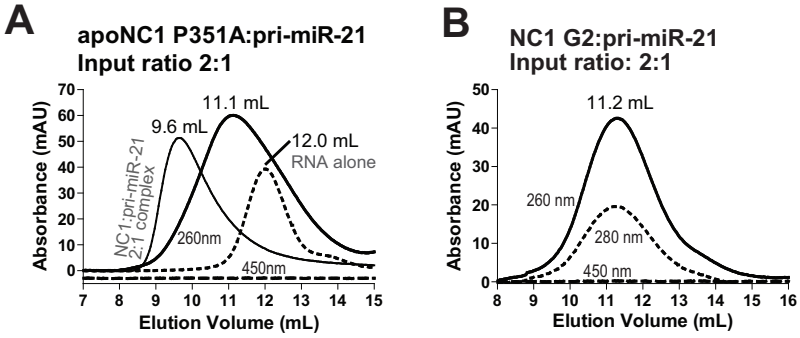


Figure S7. Size exclusion chromatograms of apoNC1 P351A in complex with pri-miR-21. Related to Figure 5

The procedure and condition are similar to those used in Figure 5. The inputs contained 2 μM of pri-miRNA and 4 μM of the dimeric protein.

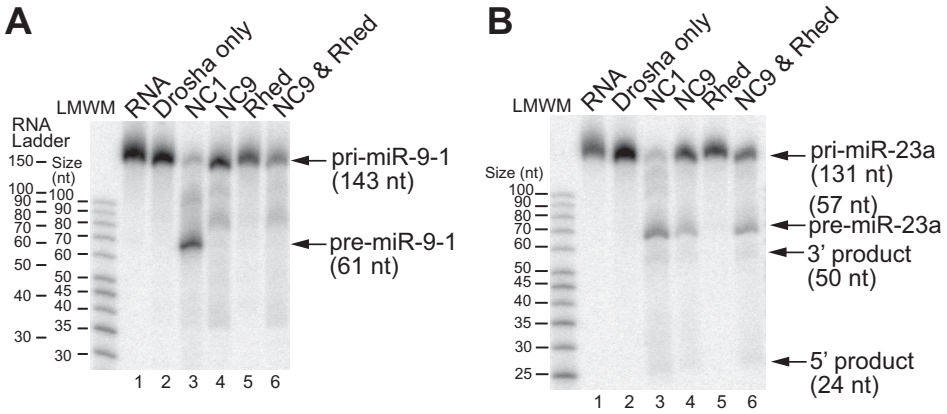


Figure S8. Reconstituted pri-miRNA processing assays indicate that the covalent linkage between the Rhed and the dsRBDs are required for pri-miRNA processing activity.

Uniformly ^{32}P-labeled pri-miRNAs were incubated with His$_6$-Drosha$^{390\text{-}1374}$ and indicated DGCR8 fragments. The concentrations of the DGCR8 proteins are 25 nM for NC1, 150 nM for NC9 and RHED. The reactions were analyzed using denaturing 15% polyacrylamide gel electrophoresis (PAGE) and autoradiography. Low molecular weight marker, LMWM. Relationship between LMWM and a true RNA ladder in 15% gels is shown in panel (**A**). Purity of RHED and NC9 is shown in Figure S1.

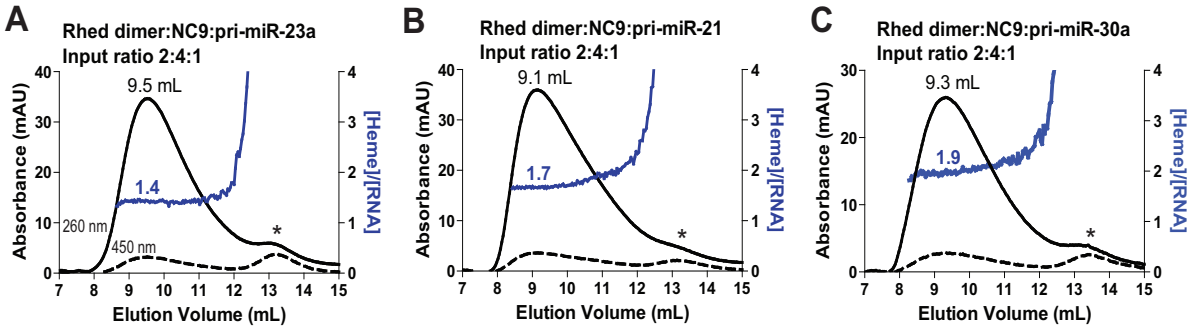


Figure S9. Size exclusion chromatograms of the Rhed and NC9 in complex with pri-miRNAs.

The procedure and condition are similar to those used in Figure 2. The inputs contained 4 μM RHED dimer, 8 μM NC9 and 2 μM pri-miRNAs. The asterisk indicates a free-RHED peak. Solid black lines indicate A_{260} and dashed lines show A_{450}. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis.

Table S1. Sequences and extinction coefficients of the pri-miRNA fragments used in the study. Related to Figures 1 and 2, and Table 1

A. pri-miRNA fragments containing all essential elements for processing

pri-miR-380 (ε_{260} = 1174 mM^{-1} cm^{-1})

GGAGAGGAAAGAGACACCGGCUCUGACCUCAGCCCUCUCCAAGGUACCUGAAAAGAUG
GUUGACCAUAGAACAUGCGCUAUCUCUGUGUCGUAUGUAAUAUGGUCCACAUCUUCUCA
AUAUCAAAUUCAGUCAUAGAGGGCUUCCC

pri-miR-9-1 (ε_{260} = 1165 mM^{-1} cm^{-1})

GGCUGCGUGGAAGAGGCGGCGACAGCAGCCAGGAGGCGGGGUUGGUUGUUAUCUUUG
GUUAUCUAGCUGUAUGAGUGGUGUGGAGUCUUCAUAAAGCUAGAUAACCGAAAGUAAAA
AUAACCCCAUACACUGCGCAGAGGGGC

pri-miR-21 (ε_{260} = 1120 mM^{-1} cm^{-1})

GGCCUACCAUCGUGACAUCUCCAUGGCUGUACCACCUUGUCGGGUAGCUUAUCAGACU
GAUGUUGACUGUUGAAUCUCAUGGCAACACCAGUCGAUGGGCUGUCUGACAUUUUGGU
AUCUUUCAUCUGACCAUCCAUAUC

pri-miR-23a (ε_{260} = 1058 mM^{-1} cm^{-1})

GGCACCCCUGUGCCACGGCCGGCUGGGGUUCCUGGGGAUGGGAUUUGCUUCCUGUCA
CAAAUCACAUUGCCAGGGAUUUCCAACCGACCCUGAGCUCUGCCACCGAGGAUGCUGC
CCGGGGACGGGGUGGC

pri-miR-30a (ε_{260} = 1200 mM^{-1} cm^{-1})

GGAAAGAAGGUAUAUUGCUGUUGACAGUGAGCGACUGUAAACAUCCUCGACUGGAAGC
UGUGAAGCCACAGAUGGGCUUUCAGUCGGAUGUUUGCAGCUGCCUACUGCCUCGGACU
UCAAGGGGCUACUUUAGGAGCAAUUAUCUUGUUU

B. Apical junction models

aj-miR-23a-C (24 bp in the stem) (ε_{260} = 520 mM^{-1} cm^{-1})

GGCUGGGGUUCCUGGGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGCCAGGGAUU
UCCAACC

aj-miR-23a-C-GAAA (24 bp in the stem)

GGCUGGGGUUCCUGGGGAUGGGAUUUGgaaaCAAAUCACAUUGCCAGGGAUUUCCAACC

aj-miR-23a-D (20 bp in the stem) (ε_{260} = 450 mM^{-1} cm^{-1})

GGGUUCCUGGGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGCCAGGGAUUUCC

aj-miR-23a-E (11 bp in the stem) (ε_{260} = 289 mM^{-1} cm^{-1})

GGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGC

aj-miR-23a-F (7 bp in the stem) (ε_{260} = 191 mM^{-1} cm^{-1})

GGAUUUGCUUCCUGUCACAAAUCC

aj-miR-21-D (18 bp in the stem) (ε_{260} = 451 mM^{-1} cm^{-1})

gGCUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGUCGAUGGGCc

aj-miR-21-E (10 bp in the stem) (ε_{260} = 314 mM^{-1} cm^{-1})

ggCUGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGcc

C. Basal junction models

bj-miR-21 (ε_{260} = 223 mM^{-1} cm^{-1})*

ggCAUGGCUGUACCACCUUGgaaaCAUUUUGGUAUCUUUCAUC

bj-miR-23a (ε_{260} = 220 mM^{-1} cm^{-1})*

ggCCCCUGUGCCACGGCCGGgaaaCCGACCCUGAGCUCUGCCA

D. ssRNA and siRNA duplex

siRNA duplex (siDGCR8-1)

sense strand 5’-CAUCGGACAAGAGUGUGAUUU-3’

anti-sense strand 3’-UUGUAGCCUGUUCUCACACUA-5’

ssRNA same as the sense strand of siDGCR8-1

All extinction coefficients were calculated using ε_{260} = M.M./(40 μg/mL), except the ones for the short RNAs marked by “*”, which were determined using alkaline hydrolysis (procedure described at http://www.scripps.edu/california/research/dna-protein-research/forms/biopolymercalc2.html)

Materials and Methods

Plasmids

Mutagenesis was performed using the 4-primer PCR method. For N-flag-DGCR8 mammalian expression plasmids, mutant DGCR8 coding sequences were inserted between BamHI and EcoRI sites in the pCMV-Tag2A vector. For bacterial expression of NC1 R322A/R325A and R341A/K342A, DGCR8-coding sequences were inserted between NdeI and EcoRI sites in the pET-24a$^+$ vector. For expression of NC1 with an N-terminal His_6 tag (G1-G6 and corresponding wild type), DGCR8 sequences were inserted between BamHI and EcoRI sites in the pRSF-Duet1 vector. For RHED-His_6 mutants, the coding sequences were inserted between NdeI and NotI in pET-24a$^+$. The transcription templates for pri-miR-30a, pri-miR-21 and pri-miR-380 have been reported (Faller et al., 2007; Faller et al., 2010; Barr et al., 2012). For pri-miR-9-1 and pri-miR-23a transcription templates, pri-miRNA sequences were amplified from human genomic DNA, and were inserted between EcoRI and PstI sites in the pUC19 vector along with a T7 promoter. The coding sequences in all plasmids were confirmed using sequencing.

Expression, Purification, and Characterization of Recombinant DGCR8 Proteins

Human NC1 (wild type and mutants; heme-bound and apo forms) and NC9 proteins were expressed and purified as previously described (Faller et al., 2007; Barr et al., 2011; Barr et al., 2012). RHED-His_6 (wild type and mutants) proteins were expressed and purified using cation exchange followed by size exclusion chromatography, same as those for NC1. Immobilized metal ion affinity chromatography was not performed to avoid the use of imidazole, which might compromise the stability of heme when incubated with the proteins for an extended period of time. The purified proteins were stored in the SEC buffer containing 20 mM Tris pH 8.0, 400 mM NaCl and 1 mM DTT, except the wild-type apoNC1 which was stored in 50 mM MOPS at pH 6.0, 400 mM NaCl and 1 mM DTT. Electronic absorption spectra were recorded at room temperature on a Cary 300 spectrophotometer with bandwidth set to 1 nm.

The ability of DGCR8 mutants to bind the nitrocellulose membrane was examined by filtering them through the membrane in the presence or absence of pri-miRNAs and blotting using anti-DGCR8 antibodies.

Transcription and Purification of pri-miRNAs

pri-miRNA fragments were produced using *in vitro* transcription and were purified using denaturing PAGE. Linearized plasmids were used as the transcription templates for pri-miRNAs. The transcription templates for apical junctions were amplified from pri-miRNA-coding plasmids using PCR. The transcription templates for basal junctions were synthesized. The RNAs were purified using denaturing PAGE and their concentrations were determined using the extinction coefficients listed in Table S1.

pri-miRNA-Binding and Processing Assays

These assays were performed as described previously (Faller et al., 2007). Briefly, for filter-binding assays, a trace amount of ^{32}P-labeled pri-miRNA was incubated with DGCR8 proteins at

room temperature for 30 min. For competition filter-binding assays, unlabeled competitor RNAs were also included in the binding reactions (Faller et al., 2010). The mixtures were filtered through nitrocellulose (EMD Millipore) and positively changed nylon (GE Healthcare) membranes. The autoradiography images of the membranes were analyzed using Quantity One (Bio-Rad version 4.4.1). The data were fit and graphed using PRISM (GraphPad version 4).

SEC analyses were performed at room temperature. The NC1 or Rhed proteins were incubated with annealed pri-miRNAs at the indicated concentrations for >5 min; these binding reactions contained a total of 233 mM NaCl. The mixtures were analyzed using an ÄKTA Purifier chromatography system and a Superdex 200 10/300 GL column (GE Healthcare), with a running buffer containing 20 mM Tris (pH 8.0) and 80 mM NaCl. After baseline subtraction, A_{450} was used to calculate DGCR8-bound heme concentration. The contribution of heme-bound DGCR8 to A_{260} was calculated based on the A_{260}/A_{450}ratio of the protein and was subtracted from the A_{260} values in the chromatograms. The remaining A_{260} was used to calculate the RNA concentration. The chromatogram plots were generated using PRISM.

For reconstituted pri-miRNA processing assays (Barr and Guo, 2014), uniformly ^{32}P-labeled pri-miRNAs were annealed and incubated at 37°C for 30 or 45 min with purified recombinant His_6-Drosha$^{390-1374}$ and DGCR8 proteins. The reactions were analyzed using 7 M urea 15% PAGE and autoradiography.

Live-Cell pri-miRNA Processing Assays

Detailed procedures have been described previously (Weitz et al., 2014). Briefly, HeLa Tet-On cells (Clontech) were transfected with reporter and/or N-flag-DGCR8-expression plasmid. Cells were immediately induced with 2 μg/ml doxycycline and imaged 18–24 hr later. Total eYFP and mCherry intensities for individual cells were fit by linear regression (y = slope * x), and slopes were obtained. p values were determined using the linear regression function of PRISM. Expression levels of N-flag-DGCR8 were analyzed using immunoblotting (Gong et al., 2012).

The mCherry-pri-miRNA fusion and the eYFP mRNA levels were determined using qRT-PCR as described (Weitz et al., 2014). The miRNA levels were measured using TaqMan assays (Life Technologies).

Author Contributions

J.Q. performed most protein-RNA-binding assays (filter binding and SEC) and mutagenesis analyses. J.P.J. did most in vitro pri-miRNA processing assays. S.H.W. performed the live-cell assays. G.S. initiated the SEC analyses and purified the RNA truncations. R.S. initiated the project. F.G., J.Q., J.P.J., and S.H.W. designed most experiments. F.G., J.Q., and J.P.J. wrote the manuscript with input from other authors.

Acknowledgments

We thank R. Johnson, D. Black, K. Yeom, and I. Barr for critical comments of the manuscript; S. Weiss for support; and Y. Chen, M. Gong, S. Griner, and N. Prakash for technical assistance. This work was partially supported by NIH grant GM080563 (to F.G.), Ruth L. Kirschstein National Research Service Award GM007185 (to J.Q. and G.S.), NIH grant T32GM008496 (to S.H.W.), and a Stein Oppenheimer Endowment Award (to F.G.).

References

Auyeung, V.C., Ulitsky, I., McGeary, S.E., and Bartel, D.P. (2013). Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell *152*, 844-858.

Axtell, M.J., Westholm, J.O., and Lai, E.C. (2011). Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome biology *12*, 221.

Barr, I., and Guo, F. (2014). Primary microRNA processing assay reconstituted using recombinant Drosha and DGCR8. Methods Mol Biol *1095*, 73-86.
Barr, I., Smith, A.T., Chen, Y., Senturia, R., Burstyn, J.N., and Guo, F. (2012). Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. Proc Natl Acad Sci USA *109*, 1919-1924.

Barr, I., Smith, A.T., Senturia, R., Chen, Y., Scheidemantle, B.D., Burstyn, J.N., and Guo, F. (2011). DiGeorge Critical Region 8 (DGCR8) is a double-cysteine-ligated heme protein. J Biol Chem *286*, 16716-16725.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Doyle, M., Jaskiewicz, L., and Filipowicz, W. (2012). Dicer proteins and their role in gene silencing pathways. In The Enzymes: Eukaryotic RNases and their partners in RNA degradation and biogenesis, Part B, F. Guo, and F. Tamanoi, eds. (Amsterdam, Netherlands, Academic Press), pp. 1-35.

Faller, M., Matsunaga, M., Yin, S., Loo, J.A., and Guo, F. (2007). Heme is involved in microRNA processing. Nat Struct Mol Biol *14*, 23-29.

Faller, M., Toso, D., Matsunaga, M., Atanasov, I., Senturia, R., Chen, Y., Zhou, Z.H., and Guo, F. (2010). DGCR8 recognizes primary transcripts of microRNAs through highly cooperative binding and formation of higher-order structures. RNA *16*, 1570-1583.

Gilles-Gonzalez, M.A., and Gonzalez, G. (2005). Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. J Inorg Biochem *99*, 1-22.

Gong, M., Chen, Y., Senturia, R., Ulgherait, M., Faller, M., and Guo, F. (2012). Caspases cleave and inhibit the microRNA processing protein DiGeorge Critical Region 8. Protein Sci *21*, 797-808.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Guo, F. (2012). Drosha and DGCR8 in microRNA biogenesis. In The Enzymes: Eukaryotic RNases and their partners in RNA degradation and biogenesis, Part B, F. Guo, and F. Tamanoi, eds. (Amsterdam, Netherlands, Elsevier Academic Press), pp. 101-121.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.

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

Heras, S.R., Macias, S., Plass, M., Fernandez, N., Cano, D., Eyras, E., Garcia-Perez, J.L., and Caceres, J.F. (2013). The Microprocessor controls the activity of mammalian retrotransposons. Nat Struct Mol Biol *20*, 1173-1181.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. Curr Biol *14*, 2162-2167.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S.*, et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012). Coordinated activities of human dicer domains in regulatory RNA processing. J Mol Biol *422*, 466-476.

Ma, H., Wu, Y., Choi, J.G., and Wu, H. (2013). Lower and upper stem-single-stranded RNA junctions together determine the Drosha cleavage site. Proc Natl Acad Sci U S A.
Macias, S., Plass, M., Stajuda, A., Michlewski, G., Eyras, E., and Caceres, J.F. (2012). DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. Nat Struct Mol Biol *19*, 760-766.

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

Marvin, K.A., Reinking, J.L., Lee, A.J., Pardee, K., Krause, H.M., and Burstyn, J.N. (2009). Nuclear receptors homo sapiens Rev-erbbeta and Drosophila melanogaster E75 are thiolate-ligated heme proteins which undergo redox-mediated ligand switching and bind CO and NO. Biochemistry *48*, 7056-7071.

Masliah, G., Barraud, P., and Allain, F.H. (2013). RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cellular and molecular life sciences : CMLS *70*, 1875-1895.

Ni, J.Q., Zhou, R., Czech, B., Liu, L.P., Holderbaum, L., Yang-Zhou, D., Shim, H.S., Tao, R., Handler, D., Karpowicz, P.*, et al.* (2011). A genome-scale shRNA resource for transgenic RNAi in Drosophila. Nat Methods *8*, 405-407.
Roth, B.M., Ishimaru, D., and Hennig, M. (2013). The core Microprocessor component DiGeorge syndrome critical region 8 (DGCR8) is a non-specific RNA-binding protein. J Biol Chem *288*, 26785-26799.

Sellier, C., Freyermuth, F., Tabet, R., Tran, T., He, F., Ruffenach, F., Alunni, V., Moine, H., Thibault, C., Page, A.*, et al.* (2013). Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell reports *3*, 869-880.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci *19*, 1354-1365.

Senturia, R., Laganowsky, A., Barr, I., Scheidemantle, B.D., and Guo, F. (2012). Dimerization and heme binding are conserved in amphibian and starfish homologues of the microRNA processing protein DGCR8. PLoS One *7*, e39688.

Shiohama, A., Sasaki, T., Noda, S., Minoshima, S., and Shimizu, N. (2007). Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins. Exp Cell Res *313*, 4196-4207.
Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R.*, et al.* (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet *37*, 1281-1288.

Sohn, S.Y., Bae, W.J., Kim, J.J., Yeom, K.H., Kim, V.N., and Cho, Y. (2007). Crystal structure of human DGCR8 core. Nat Struct Mol Biol *14*, 847-853.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nat Struct Mol Biol *18*, 1153-1158.

Weitz, S.H., Gong, M., Barr, I., Weiss, S., and Guo, F. (2014). Processing of microRNA primary transcripts requires heme in mammalian cells. Proc Natl Acad Sci U S A *111*, 1861-1866.

Yeom, K.H., Lee, Y., Han, J., Suh, M.R., and Kim, V.N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Res *34*, 4622-4629.

Yin, L., Wu, N., Curtin, J.C., Qatanani, M., Szwergold, N.R., Reid, R.A., Waitt, G.M., Parks, D.J., Pearce, K.H., Wisely, G.B.*, et al.* (2007). Rev-erbalpha, a heme sensor that coordinates metabolic and circadian pathways. Science *318*, 1786-1789.

Zeng, Y., and Cullen, B.R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J Biol Chem *280*, 27595-27603.
Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J *24*, 138-148.

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

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.

Chapter 3: An Fe(III) heme pool regulates the Microprocessor Complex

This chapter is based on a collaboration between myself and Sara Weitz. I discovered a group mutation in the C-terminal region of the Rhed that did not bind heme. I characterized the group mutant as well as the single mutations in isolation, measuring RNA binding and heme-binding properties. Using these mutants, and some unpublished work by Ian Barr on the P351A mutation, were able to calculate a threshold concentration of heme that was necessary to activate pri-miRNA processing. Sara performed the cellular assays and Sara Weitz and I wrote the manuscript with assistance from Dr. Guo. This work will be submitted for publication.

Introduction

Heme is an essential cofactor for life. It participates in numerous biological pathways, by serving as a prosthetic group for heme proteins, or as a signaling molecule to regulate cellular processes (Ponka, 1999; Tracz et al., 2007; Zhang, 2011). For cells to function normally, a fine balance is maintained between heme synthesis and degradation through feedback and feedforward mechanisms (Khan and Quigley, 2011; Ryter and Tyrrell, 2000; Ye and Zhang, 2004). Over-abundance of heme causes severe cell and tissue damages. Conversely, heme deficiency results in anemia or porphyria. Most heme molecules in cells and tissues are already bound to their host proteins. Therefore, it is the pool of labile heme that determines whether newly synthesized heme proteins are loaded, and whether a signaling pathway is turned on or off. Labile heme in cells is transient and is dynamically regulated, leading to technical challenges in measuring its concentration (Sassa, 2004) (Hanna and Reddi, 2016). Based on characterization of the heme degradation enzymes, heme oxygenases (K_M = ~1 μM), it was proposed that the concentration of available heme may be just below the K_M, probably around 100 nM. Another theory suggests that available heme molecules are bound to heme transporter proteins and there is little to no free heme in cells (Severance and Hamza, 2009).

Heme activates miRNA maturation by binding to DGCR8 (*D*i*G*eorge *c*ritical *r*egion gene *8*) (Barr et al., 2012; Faller et al., 2007; Weitz et al., 2014). DGCR8 and the ribonuclease III Drosha form the Microprocessor complex (MC), which recognizes and cleave pri-miRNAs to produce the processing intermediates, precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004). DGCR8 contains an RNA-binding heme domain (Rhed, residues 276-498 in humans) that binds to the ends of pri-miRNA hairpins and works with the two double-stranded RNA-binding domains (dsRBDs) of DGCR8 to achieve high pri-miRNA-binding affinity and specificity (Figure 1A) (Quick-Cleveland et al., 2014). At steady state, the heme iron may adopt either 2+ or 3+ oxidation state. Only the Fe(III) heme availability directly affects pri-miRNA processing, as DGCR8 strongly prefers binding Fe(III) heme over Fe(II) heme and only Fe(III) heme can activate DGCR8 (Barr et al., 2012; Barr et al., 2011). DGCR8 binds heme as a dimer using Cys352 from each subunit as the coaxial ligands (Barr et al., 2011; Faller et al., 2007; Senturia et al., 2010). The electron-rich dual cysteine ligation is key to the extraordinary specificity of DGCR8 for Fe(III) heme.

The threshold concentration of Fe(III) heme necessary to activate pri-miRNA processing may be determined by characterizing DGCR8 mutants with various affinities for Fe(III) heme. The intracellular Fe(III) heme availability should be between the dissociation constants (K_d) of an active DGCR8 with the lowest Fe(III) heme affinity and an inactive DGCR8 mutant with the highest Fe(III) heme affinity. Cys352 and the surrounding residues, including Ile350, Pro351, and Leu353, generate the “IPCL” motif that is critical for heme binding. Mutation of any single one of these residues to alanine abolishes DGCR8 activity in cells (Weitz et al., 2014). Out of these mutants, only P351A is still capable of binding Fe(III) heme in vitro, though with much reduced affinity (Barr et al., 2011). The heme-binding affinities of heme proteins may be compared by measuring their dissociation rates (k_{off}) (Hargrove et al., 1996). The k_{off} of P351A-heme complex is around 5×10^{-4} s^{-1} at room temperature, whereas the k_{off} of wild-type DGCR8 is very slow with no Fe(III) heme dissociation observed in 4 days ($k_{off} << 3 \times 10^{-6}$ s^{-1}) (Barr et al., 2011). The large gap between the

Fe(III) heme affinities of the active wild-type DGCR8 and the inactive P351A mutant leaves it uncertain what the available Fe(III) heme concentration is in cells.

Here we identify DGCR8 mutants that have reduced affinities for Fe(III) heme but maintain their activity. In combination with the inactive mutants described above, our results suggest that the Fe(III) heme available to activate the MC in HeLa cells, falls in a surprisingly narrow and low concentration range. We validated this estimate by manipulating the cellular heme environments and measuring MC activity. These findings support the idea that this is a pool of Fe(III) heme that regulates pri-miRNA processing in cells. This has important implications as to how Fe(III) heme is produced and transported in mammalian cells.

Results

The DGCR8 G2 mutant does not bind heme

Through systematic mutagenesis of basic residues within the C-terminal region of Rhed, designed for the purpose of another study (Quick-Cleveland et al., 2014), we identified a group mutant G2 (containing R441A, K446A and R447A) that is severely defective in heme binding. The mutations were introduced to NC1, a truncated construct of DGCR8 (a.a. 276-751) that includes all functionally important regions for pri-miRNA processing, namely the Rhed, the dsRBDs and the C-terminal tail (CTT) (Figure 1A). We overexpressed the mutants in *E. coli*, purified them to homogeneity and examined the presence of heme using electronic absorption spectroscopy. No intense absorption was observed in the 400-500 nm region (Figure 1B). The minor residual Söret peak was mostly at 424 nm, not the 450 nm characteristic for Fe(III) heme ligated by two cysteine residues of DGCR8 (Figure 1B inset). These results indicate that one or more of the residues mutated in G2 are necessary for DGCR8-heme association.

Next, we measured the activity of the DGCR8 G2 mutant using *in vitro* and cellular assays. In reconstituted pri-miRNA processing assays, we incubated ^{32}P-labeled pri-miRNAs with purified NC1 G2 and His_6-Drosha$^{390-1374}$ proteins. Denaturing polyacrylamide gel analyses indicated that

the pri-miRNA processing activity of G2 was abolished (Figure 1C,D). To test the cellular activity of G2, we used a fluorescent-based live-cell pri-miRNA processing reporter assay (Weitz et al., 2014). In this assay, a pri-miRNA (pri-miR-9-1) is inserted into the 3′UTR of mCherry so that processing of the pri-miRNA results in degradation of the fusion mRNA and thereby a reduction of mCherry fluorescence. The reporter plasmid simultaneously transcribes *eYFP* and the *mCherry-pri-miRNA* fusion through the use of a bi-directional inducible promoter. The eYFP and mCherry fluorescence signals are measured for individual cells and the slope of eYFP vs mCherry indicates the cellular pri-miRNA processing efficiency. Co-transfection of the reporter with the expression plasmid for full-length DGCR8 (N-flag-DGCR8) increases pri-miRNA processing efficiency, with the fluorescent slope increased by ~2 fold in HeLa cells cultured in complete medium (Figure 1E). This assay allows us to measure the activity of DGCR8 mutants in cells. Because endogenous DGCR8 is expressed at very low levels in HeLa cells, changes in fluorescence slope mostly reflect the activity of ectopically expressed N-flag-DGCR8 proteins. All inactive DGCR8 mutants characterized in our previous studies give fluorescence slopes similar to the vector control (Quick-Cleveland et al., 2014; Weitz et al., 2014). Here we found that expression of N-flag-DGCR8 G2 in HeLa cell cultured in complete medium resulted in a fluorescent slope close to that of the vector control and much lower than wild-type DGCR8 (Figure 1E). Quantification of mature miR-9 produced from the reporter supported the conclusion that G2 is inactive in cells (Figure 1F). Western blot analysis showed that G2's lack of activity was not due to reduced protein expression relative to the wild type (Figure 1G). Altogether, our results suggest that the G2 mutations cause DGCR8 to lose Fe(III) heme binding and thereby the pri-miRNA processing activity.

G2 single mutants show reduced affinity for heme, but are active in pri-miRNA processing

To determine which residue is responsible for the heme-binding defect of G2, we engineered the individual mutations in the context of NC1 and tested their ability to bind heme

when expressed in *E. coli*. As Lys446 and Arg447 are immediately next to each other, we also tested if these two residues function together by characterizing the double mutant K446A/R447A. Similar to G2, the purified NC1 K446A/R447A protein did not contain Fe(III) heme (Figure S1), suggesting that simultaneous mutation of these residues to a large extent accounts for the heme-binding defect. Somewhat surprisingly, NC1 R441A, K446A, and R447A single mutants were all capable of binding Fe(III) heme (Figure 2A-C). These results suggest that Lys446 and Arg447 serve a redundant function in heme binding. These basic residues may stabilize heme by directly contacting one or both propionate groups, as has been shown in other heme proteins such as myoglobin, human catalase I and hemopexin (Smith et al., 2010).

We then examined whether the single mutants have reduced heme affinity by measuring their rates of heme dissociation. We incubated the mutants with a five-fold excess of apomyoglobin, which has a high affinity for heme and serves as a heme scavenger. The Fe(III) heme dissociation from DGCR8 is indicated by a reduction in the 450-nm Söret peak with a concurrent appearance of a 409-nm peak indicative of Fe(III) heme transferring to myoglobin. Among the three NC1 mutants, R441A had the strongest association with Fe(III) heme, with the half-life of heme loss ($t_{1/2}$) >8 hr (k_{off} = 2.3 × 10^{-5} s^{-1}) (Figure 2A). K446A lost its heme faster with $t_{1/2}$ ≈ 3.3 hr (k_{off} = 8.7 × 10^{-5} s^{-1}) (Figure 2B). The dissociation of heme from NC1-R447A was very slow in the initial 4 hr, but subsequently sped up with most of the Fe(III) heme transferred by the 8th hr (Figure 2C). The R447A data could not be fit using a single exponential function, suggesting that the mutant loses heme through a multi-step process. Compared to the previously characterized heme-binding mutations, the three G2 single mutations more modestly reduce the affinity for Fe(III) heme.

We subsequently engineered the G2 double and single mutations in the context of N-flag-DGCR8 and measured their cellular activity. In HeLa cells cultured in the complete medium, K446A/R447A is inactive in pri-miRNA processing, whereas the three single mutants were as active as wild-type N-flag-DGCR8, as indicated by both fluorescence slopes and mature miR-9

production (Figure 2D,E). Western blots showed that the mutant proteins express well in HeLa cells (Figure 2F), ruling out the possibility that the lack of activity for K446A/R447A was caused by reduced expression. Furthermore, similar fluorescence slopes and miR-9 expression levels were obtained when the mutants were expressed at levels higher than or similar to the wild type (the latter data not shown). Therefore, the Fe(III) heme affinity of the G2 single mutants must be sufficient to acquire heme and to process pri-miRNAs under this experimental condition.

K446A and P351A provide an estimate of intracellular Fe(III) heme availability

Among the three G2 single mutants, K446A has the lowest affinity for Fe(III) heme. The k_{off} of NC1-K446A is only five-fold slower than that of P351A, the inactive mutant with the highest affinity for Fe(III) heme. Assuming this dramatic change in activity is caused solely by differences in Fe(III) heme affinities, this pair of DGCR8 mutants allow the intracellular available Fe(III) heme concentration to be estimated with great precision.

To determine the affinity of K446A and P351A for Fe(III) heme, we attempted to measure their Fe(III) heme association rate k_{on} (as $K_d = k_{off} / k_{on}$). We previously reported that purified heme-free (apo) NC1-P351A dimer can bind Fe(III) heme to reconstitute a complex with an absorption spectrum very similar to that of the native complex (Barr et al., 2011). In these measurements, we monitored over time the growth of the Söret peak at 447 nm as an indicator of complex formation. This binding reaction went to completion quickly, so we employed the stop flow method. The data were best fit to a two-phase association curve, giving two pseudo-first-order rates k_1 and k_2 (Figure S3A). To obtain the rate for this second-order reaction, we measured the association rates at four apoNC1-P351A concentrations, all in large excess of the Fe(III) heme concentration (Figure S3B). The pseudo-first-order rates should have a linear relationship with the protein concentration and the slope gives k_{on}. However, both k_1 and k_2 remained roughly unchanged over the apoNC1-P351A concentration range. This is most likely because Fe(III) heme tends to stack in aqueous solutions, forming a variety of oligomers. Unstacking of heme from oligomers may be

the rate-limiting step of the binding reaction. Imidazole can unstack Fe(III) heme to produce monomers, but it appeared to interfere with the DGCR8-Fe(III)-heme interaction. It is currently not possible to measure the second-order k_{on} for DGCR8 and monomeric Fe(III) heme. Nevertheless, our measurements indicate that association between DGCR8 and Fe(III) heme is very fast.

Previous characterization of many heme proteins revealed that their k_{on}'s for heme are very similar and thus the K_d values are largely determined by the k_{off} (Bhakta and Wilks, 2006; Gaudin et al., 2011; Hargrove et al., 1996; Liu et al., 2008; Nygaard et al., 2006; Owens et al., 2012). The k_{on}'s for 35 globins, BSA, and bacterial heme uptake proteins IsdA and Rv0203 are all about 1×10^8 M^{-1} s^{-1}. Assuming DGCR8 binds Fe(III) heme with a similar k_{on}, we estimate that the K_d of P351A for Fe(III) heme is ~5 pM, whereas that for K446A is ~1 pM. The concentration range from 1 - 5 pM defines a threshold of the Fe(III) heme concentration available to support pri-miRNA processing within HeLa cells.

Validation of the Fe(III) heme availability model

Our estimated range of intracellular Fe(III) heme availability rests on the assumption that the differing affinities of K446A and P351A for Fe(III) heme is responsible for the observed change in activity. However, it is possible that mutational effects unrelated to heme binding are involved. We therefore tested our model by monitoring the pri-miRNA processing activity of these DGCR8 mutants while altering heme levels in cells. We used four different heme conditions, including (1) the complete medium (CM), (2) complete medium containing 1 mM heme synthesis inhibitor succinylacetone (CM+SA), (3) heme-depleted medium (HDM) in which the heme in serum is partially (~50%) removed, and (4) heme-depleted media containing 1 mM SA (HDM+SA). The pri-miRNA processing activity of wild-type N-flag-DGCR8 was not sensitive to a reduced amount of heme in the medium or inhibition of endogenous heme synthesis alone (Figure 3A). However, a more severe heme deficiency, caused by combination of the two treatments, rendered the activity of N-flag-DGCR8 greatly reduced (Figure 3A). Measurements of mature miR-9 produced

were generally consistent with the fluorescence slopes, although the larger errors limited further insight (Figure 3B). We were able to rescue the pri-miRNA processing defect under the HDM+SA condition by adding increasing concentrations of hemin to the media (Figure 3C). A full rescue was achieved at 1 and 10 μM hemin. Even 0.2 μM hemin was able to slightly but significantly rescue the defect. The dose-dependent rescue excluded the possibility that the reduced N-flag-DGCR8 activity is an unintended effect of depleting heme from the serum or the heme biosynthesis inhibitor. Anti-DGCR8 immunoblotting ruled out another possibility that reduced N-flag-DGCR8 expression level is responsible for the pri-miRNA processing deficiency (Figure 3D). Overall, these results suggest that, under typical culture conditions, Fe(III) heme is not limiting pri-miRNA processing in HeLa cells. Similar effects, but to much less extents, have been observed in HeLa cells without ectopic expression of N-flag-DGCR8 (Weitz et al., 2014). Co-expression of N-flag-DGCR8 in HeLa cells provides a sensitive system for testing the effect of Fe(III) heme availability on the activity of DGCR8 mutants.

In our model, the K_d of K446A for Fe(III) heme is close to the concentration of available Fe(III) heme in HeLa cells cultured in complete media. Therefore, the activity of this mutant should be more sensitive to heme depletion than the wild type. Indeed, in the cellular assays for N-flag-DGCR8 K446A, even the heme-depleted medium alone caused a large reduction of the fluorescence slope, from 2.21 ± 0.07 in CM to 1.67 ± 0.07 in HDM (Figure 4A). Combination of HDM and SA treatments further deactivated K446A, with the fluorescence slope decreased to 1.15 ± 0.05, also lower than that (1.31 ± 0.06) of the wild type (Figure 4A). These changes are likely to be direct effects of altering Fe(III) availability, as addition of hemin rescued the K446A activity in a dose-dependent manner.

In a parallel experiment, we tested the N-flag-DGCR8 P351A mutant. Predictably, P351A remained largely inactive under all conditions tested (Figure 4A). We tried to rescue the P351A defect by titrating in 0.2, 1, and 10 μM hemin. Indeed, when 10 μM hemin was added to the medium, the fluorescence slope for P351A increased modestly from 0.99 ± 0.04 (HDM+SA) to

1.23 ± 0.02 and this increase was statistically significant ($p < 0.001$). It is likely that the presence of 10 μM hemin in the medium raised the intracellular Fe(III) heme levels so that P351A was able to partially or transiently associate with heme.

In addition to reducing affinity for Fe(III) heme, the P351A mutation may inactivate DGCR8 in cells by other mechanisms, i.e. by affecting the Fe(III) heme-induced conformational change. We found that purified apoNC1-P351A is activated for pri-miRNA processing by Fe(III) heme (Figure S4) in a way very similar to wild-type apoNC1 (Barr et al., 2012), thereby ruling out this possibility. Overall, this series of validation experiments strongly supports our model that the intracellular Fe(III) heme availability in HeLa cells under typical culture conditions is between the K_d's of the K446A and P351A mutants for Fe(III) heme.

Discussion

The presence of an Fe(III) heme pool that sharply determines miRNA maturation efficiency indicates cross-regulation between heme and miRNA pathways. Heme is synthesized in the Fe(II) form (Dailey, 2002) and has to be oxidized to the Fe(III) state when activating miRNA maturation. The heme iron oxidization must occur during trafficking from mitochondria, the site of synthesis, to its destination, which could be the cytoplasm where DGCR8 is translated, or the nucleus where pri-miRNA processing occurs. The interconversion between the Fe(II) and Fe(III) forms is presumably regulated by the redox environment of the cell, including metabolic rates, cell growth and oxygen availability. Redox regulation of available heme may also be linked to heme degradation, and import and export across cellular membranes (Khan and Quigley, 2011; Severance and Hamza, 2009).

The ultra-low effective concentration (1-5 pM) of available Fe(III) heme seems to present a paradox. Considering that the volume of a typical HeLa cell is ~2,000 μm^3, this concentration means that only a few Fe(III) heme molecules exist within each cell cultured in the complete medium. This low concentration estimate agrees well with recent work from the Reddi group that

measures the labile heme concentration with a FRET-based heme sensor in yeast (Hanna and Reddi, 2016). Based on the low and high affinity sensors HS1-M7A and HS1, they estimate fewer than six Fe(II) heme molecules in the nucleus, and report a total cell Fe(II) heme concentration of ~ 1uM. This low heme concentration leaves little room for further reduction under heme deficient conditions. Additionally, it may be challenging for these few Fe(III) heme molecules to "find" the apo DGCR8 proteins whether they are in the nucleus or cytoplasm. This paradox may be resolved by the possibility that a currently unknown Fe(III) heme-specific transporter (or chaperone) loads and activates DGCR8. In this scenario, the effective available Fe(III) heme concentration we observed may reflect an affinity requirement for a DGCR8 protein to receive Fe(III) heme from the transporter. The actual available Fe(III) concentration may be higher and heme deficient conditions result in fewer transporter molecules loaded with Fe(III) heme. The residues in the G2 mutant, R441A, K446A, and R447A, and a recently identified heme-free mutant, F448L could be region of the protein important for Fe(III) heme transfer to DGCR8.

Copper is another essential, tightly regulated trace element essential for biological processes. Like heme, there is very little "free" copper because of the potential for toxicity. Metallochaperones such as ATX1 in yeast bind and transfer copper to its target copper transport protein CCC2 (O'Halloran and Cullota 2000). Another example of this is the antioxidant SOD1, which requires the copper chaperone CCS to for activation in yeast (Rae and O'Halloran 1999). Although few heme methallochaperones are known, as structural study of hGRX5 showed this protein transfers Fe-S clusters to hISCA1 (Bianci and Del Conte 2014).

The function of the tightly controlled pool of available Fe(III) heme is not limited to miRNA maturation. A class of heme proteins that have been proposed to be heme sensors also used electron-rich cysteine side chain thiolate for axial ligation, and thus generally prefers Fe(III) heme over Fe(II) heme (Shimizu, 2011). These sensors are involved in circadian rhythm, transcription regulation of metabolic pathways, heme catabolism, ion channel activities and protein degradation.

The pool of Fe(III) heme that regulates miRNA maturation may be a common link that coordinates miRNA production with other pathways important for cell physiology.

Materials and Methods

Expression, purification and characterization of DGCR8 (NC1) proteins

Expression, purification and reconstituted pri-miRNA processing assays were done as described (Barr and Guo, 2014). Proteins were expressed in the presence of 1 mM δ-ALA. The proteins were stored in 20 mM Tris pH 8.0, 400 mM NaCl and 1 mM dithiothreitol. Heme off-rate measurements were performed as previously described (Barr et al., 2011). Details are described in **Supplemental Information**.

Cellular pri-miRNA processing assays

The assay has been described previously (Weitz et al., 2014). Succinylacetone and hemin stock solutions were prepared in water and DMSO, respectively. Cells were grown either in heme-depleted media or complete media and then split into glass-bottom plates for imaging. Cells were treated for 2 hr with 1 mM SA either alone or with hemin. Cells were cotransfected with the pri-miR-9-1 reporter and the indicated N-flag-DGCR8 construct. We used twice the amount of N-flag-DGCR8 mutant plasmid as the wild type in the cotransfections. Doxycycline was added to a final concentration of 1 μM to induce reporter expression. Cells were imaged 20-24 hr post-induction.

Author Contributions

S.H.W. performed all cellular experiments. J.Q. performed all biochemical experiments except the in vitro pri-miRNA processing assays (done by J.P.J. and R.S.) and the stop-flow heme association measurements (by I.B.). All authors designed experiments. S.H.W., J.Q. and F.G. wrote the paper.

Acknowledgments

This work was supported by NIH grant GM080563 to F.G., Ruth L. Kirschstein National Research Service Award GM007185 to J.Q. and G.S., a Stein Oppenheimer Endowment Award to F.G., and a UCLA Dissertation Year Fellowship to S.H.W. The authors declare no conflict of interest.

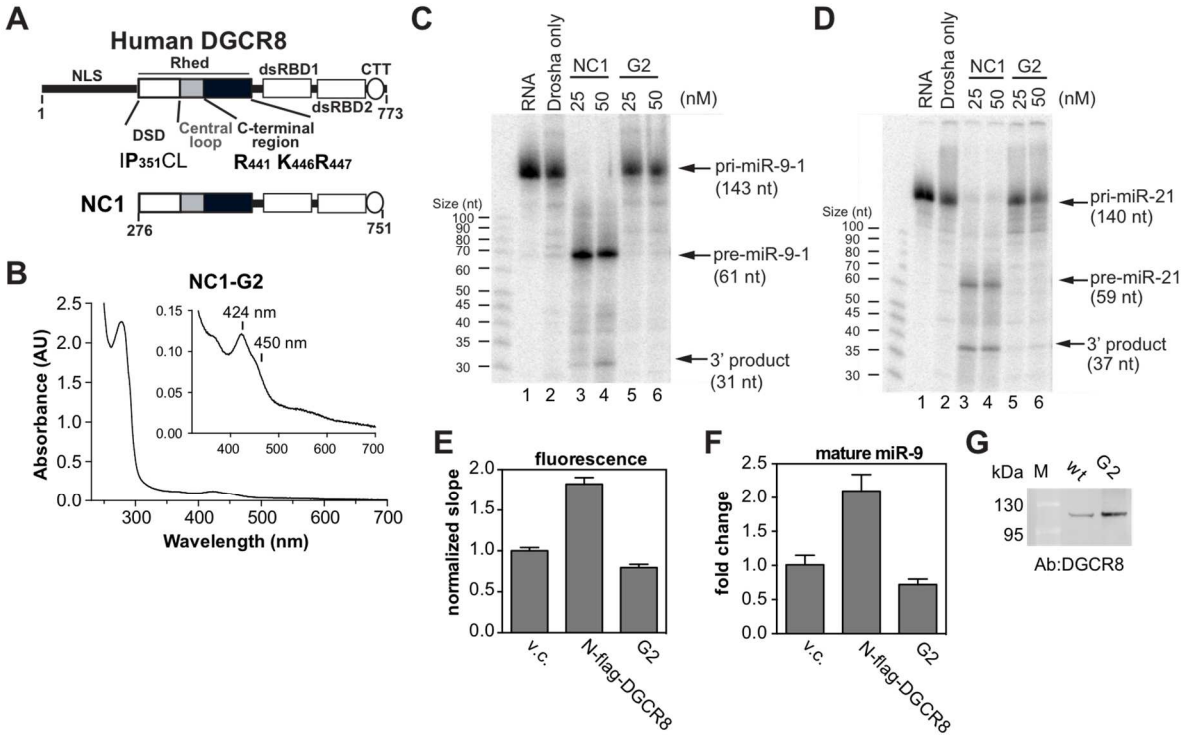


Figure 1. The DGCR8 G2 mutant fails to bind heme and is inactive in pri-miRNA processing. (A) Domain structure of human DGCR8, with relevant residues labeled. The NC1 construct is indicated underneath. (B) Electronic absorption spectrum of purified NC1 G2 with an inset showing a zoom-in. No intense 450 nm Söret peak was observed. (C,D) pri-miRNA processing assays reconstituted using purified Drosha and 25 nM NC1 dimers. (E) HeLa cells cultured in complete medium were transfected with pri-miR-9-1 reporter plus either pCMV-tag2a (vector control, v.c.) or N-flag-DGCR8 expression plasmids as indicated. Normalized fluorescence slopes are plotted. Error bars represent 95% confidence interval from

the linear fit. (F) The abundance of mature miR-9 (mean ± SD, n = 3) from (E) measured using qRT-PCR. (G) An anti-DGCR8 immunoblot with equal amounts of nuclear extracts loaded. See also Figure S1.

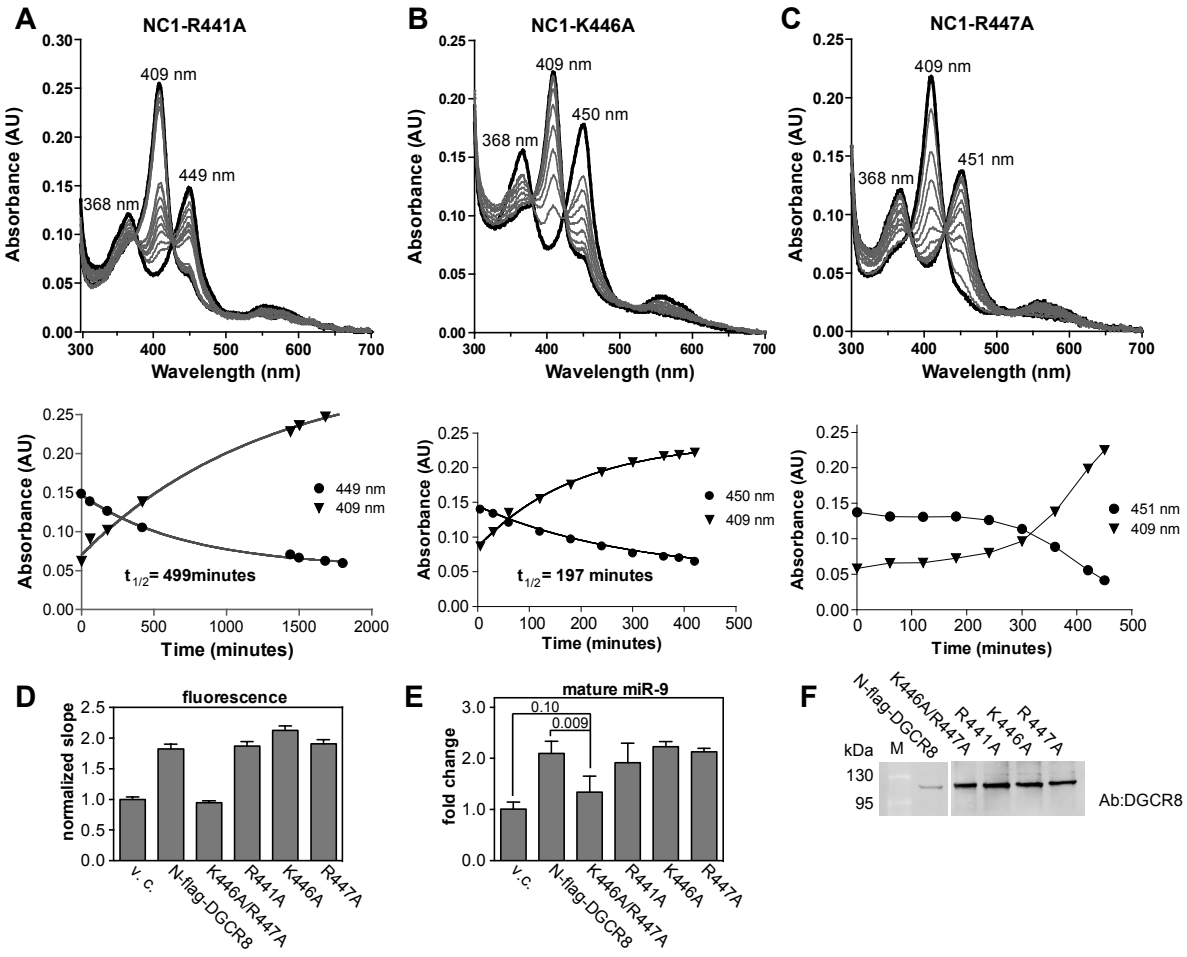


Figure 2. K446A and R447A modestly reduce affinity for Fe(III) heme but remain active in pri-miRNA processing. Fe(III) heme bound NC1 R441A (A) K446A (B) or R447A (C) was incubated with 6-fold excess of apomyoglobin. Top: electronic absorption spectra obtained at different time points. Bottom: A_{409} (aquametmyoglobin) and A_{450} (Fe(III) heme-bound NC1) time points were fit to a single exponential (A, B) or simply connected by lines (C). (D) Cellular pri-miR-9-1 processing assays in complete culture media. Error bars represent 95% CI from the linear fit. (E) Abundance of mature miR-9 (mean ± SD, n = 3). (F) An anti-NC1 immunoblot with equal amounts of nuclear extracts loaded. See also Figures S1 and S2.

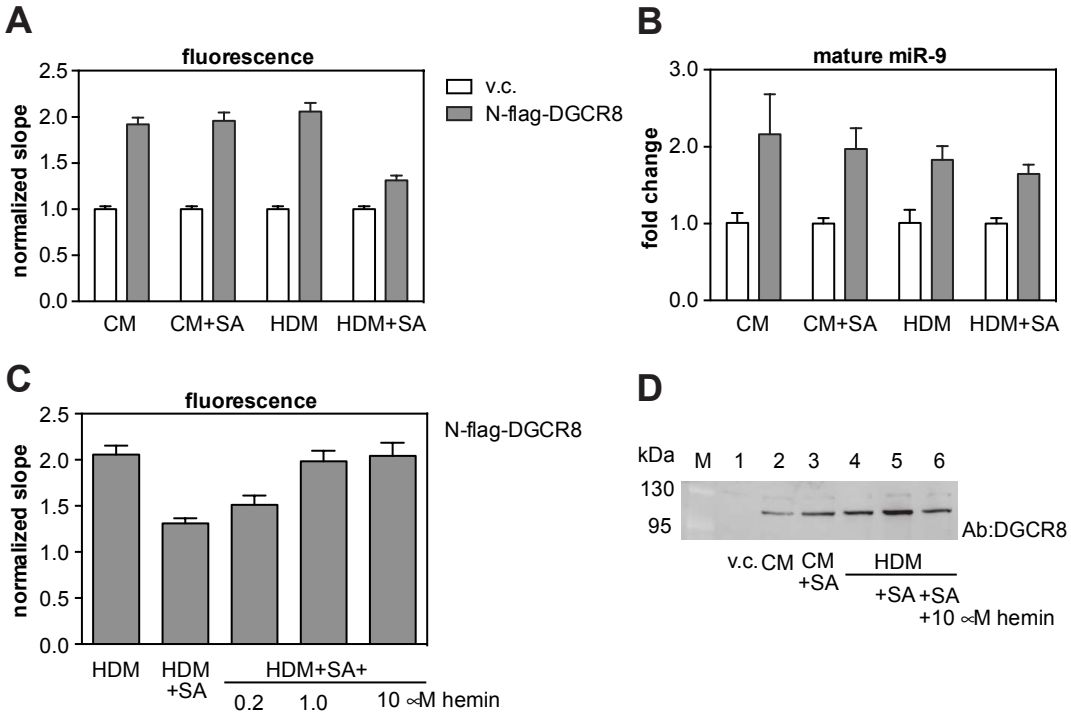


Figure 3. Only severe heme deficiency influences pri-miRNA processing efficiency of wild-type N-flag-DGCR8 in HeLa cells. Cellular pri-miR-9-1 processing assays were performed with the N-flag-DGCR8 expression plasmid cotransfected with the reporter. CM: complete media, CM+SA: complete media plus 1 mM succinylacetone, HDM: heme depleted media, HDM+SA: heme depleted media plus 1 mM succinylacetone. (A) Normalized fluorescence slopes. Error bars represent 95% CI from the linear fit. (B) Abundance of mature miR-9 (mean ± SD, n = 3). (C) Cellular pri-miR-9-1 processing assays in which HDM+SA treatment was rescued by adding 0.2, 1, or 10 μM hemin to the media. Error bars represent 95% CI from the linear fit. (D) Anti-DGCR8 immunoblots with equal amounts of nuclear extracts loaded.

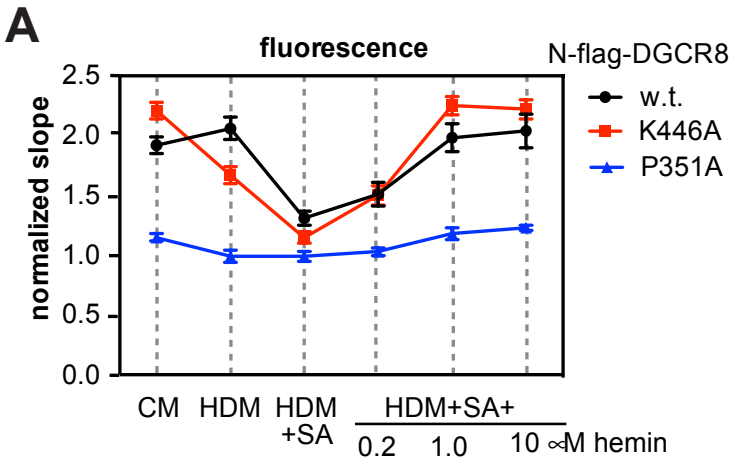


Figure 4. Validation under various heme conditions supports our model of estimating the boundaries of an Fe(III) heme pool.

(A) Cellular pri-miR-9-1 processing assay for N-flag-DGCR8 wild-type, P351A and K446A under varying heme conditions as described in Figure 3. See also Figures S3 and S4.

Supplemental Figures

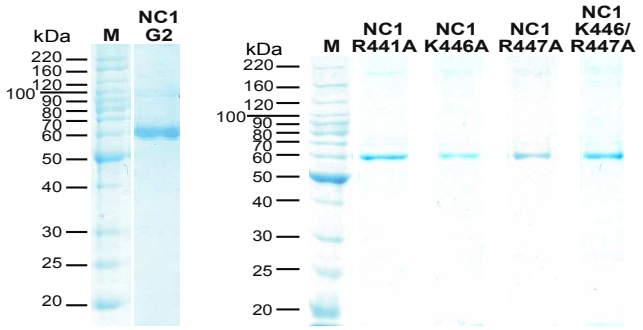


Figure S1. Coomasie-stained SDS-PAGE of purified recombinant DGCR8 proteins used in the study.

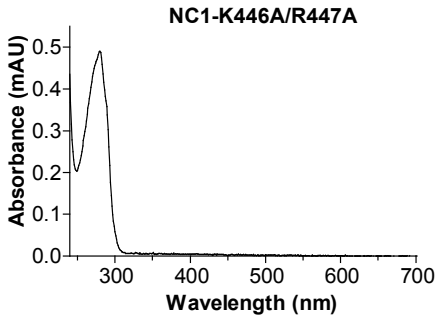


Figure S2. Electronic-absorption spectrum of NC1 K446A/R447A DCR8 mutant.

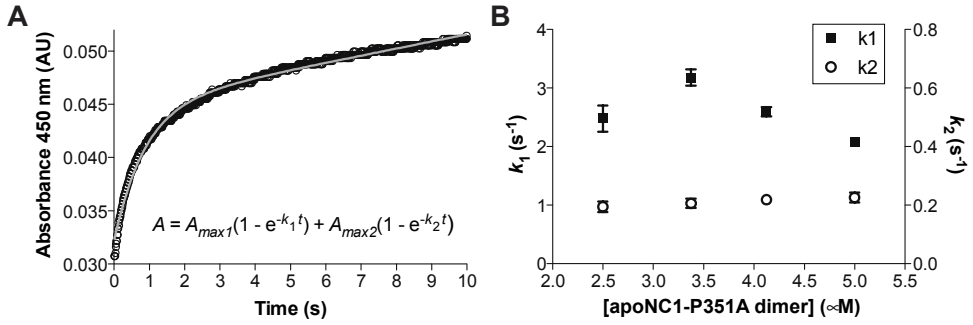


Figure S3. apoNCI-P351A heme association by stop-flow method. A) Data is fit to a two-phase association curve giving two pseudo-first order heme association rates. B) Second-order rates measurement

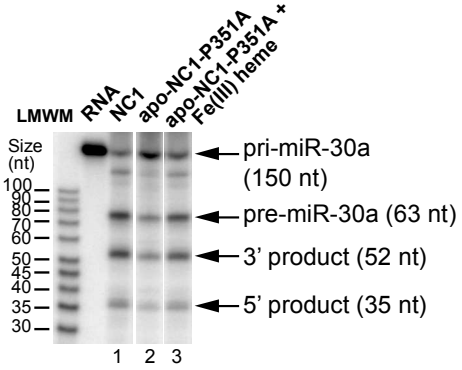


Figure S4. Apo-NC1-P351A is activated for pri-miRNA processing by Fe(III) heme similar to wild-type apo-NCI protein.

References

Barr, I., and Guo, F. (2014). Primary microRNA processing assay reconstituted using recombinant Drosha and DGCR8. Methods Mol. Biol. *1095*, 73-86.

Barr, I., Smith, A.T., Chen, Y., Senturia, R., Burstyn, J.N., and Guo, F. (2012). Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. Proc. Natl. Acad. Sci. U.S.A. *109*, 1919-1924.

Barr, I., Smith, A.T., Senturia, R., Chen, Y., Scheidemantle, B.D., Burstyn, J.N., and Guo, F. (2011). DiGeorge Critical Region 8 (DGCR8) is a double-cysteine-ligated heme protein. J. Biol. Chem. *286*, 16716-16725.

Bhakta, M.N., and Wilks, A. (2006). The mechanism of heme transfer from the cytoplasmic heme binding protein PhuS to the delta-regioselective heme oxygenase of Pseudomonas aeruginosa. Biochemistry *45*, 11642-11649.

Dailey, H.A. (2002). Terminal steps of haem biosynthesis. Biochem. Soc. Trans. *30*, 590-595.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Faller, M., Matsunaga, M., Yin, S., Loo, J.A., and Guo, F. (2007). Heme is involved in microRNA processing. Nat. Struct. Mol. Biol. *14*, 23-29.

Gaudin, C.F., Grigg, J.C., Arrieta, A.L., and Murphy, M.E. (2011). Unique heme-iron coordination by the hemoglobin receptor IsdB of Staphylococcus aureus. Biochemistry *50*, 5443-5452.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. *18*, 3016-3027.

Hargrove, M.S., Barrick, D., and Olson, J.S. (1996). The association rate constant for heme binding to globin is independent of protein structure. Biochemistry *35*, 11293-11299.

Khan, A.A., and Quigley, J.G. (2011). Control of intracellular heme levels: heme transporters and heme oxygenases. Biochim. Biophys. Acta *1813*, 668-682.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. Curr. Biol. *14*, 2162-2167.

Liu, M., Tanaka, W.N., Zhu, H., Xie, G., Dooley, D.M., and Lei, B. (2008). Direct hemin transfer from IsdA to IsdC in the iron-regulated surface determinant (Isd) heme acquisition system of Staphylococcus aureus. J. Biol. Chem. *283*, 6668-6676.

Nygaard, T.K., Blouin, G.C., Liu, M., Fukumura, M., Olson, J.S., Fabian, M., Dooley, D.M., and Lei, B. (2006). The mechanism of direct heme transfer from the streptococcal cell surface protein Shp to HtsA of the HtsABC transporter. J. Biol. Chem. *281*, 20761-20771.

Owens, C.P., Du, J., Dawson, J.H., and Goulding, C.W. (2012). Characterization of heme ligation properties of Rv0203, a secreted heme binding protein involved in Mycobacterium tuberculosis heme uptake. Biochemistry *51*, 1518-1531.

Ponka, P. (1999). Cell biology of heme. Am. J. Med. Sci. *318*, 241-256.

Quick-Cleveland, J., Jacob, J.P., Weitz, S.H., Shoffner, G., Senturia, R., and Guo, F. (2014). The DGCR8 RNA-binding heme domain recognizes primary microRNAs by clamping the Hairpin. Cell Rep. *7*, 1994-2005.

Ryter, S.W., and Tyrrell, R.M. (2000). The heme synthesis and degradation pathways: role in oxidant sensitivity. Heme oxygenase has both pro- and antioxidant properties. Free Radic Biol Med *28*, 289-309.

Sassa, S. (2004). Why heme needs to be degraded to iron, biliverdin IXalpha, and carbon monoxide? Antioxid Redox Signal *6*, 819-824.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. *19*, 1354-1365.

Severance, S., and Hamza, I. (2009). Trafficking of heme and porphyrins in metazoa. Chem. Rev. *109*, 4596-4616.

Shimizu, T. (2011). Binding of cysteine thiolate to the Fe(III) heme complex is critical for the function of heme sensor proteins. J. Inorg. Biochem.

Smith, L.J., Kahraman, A., and Thornton, J.M. (2010). Heme proteins--diversity in structural characteristics, function, and folding. Proteins *78*, 2349-2368.

Tracz, M.J., Alam, J., and Nath, K.A. (2007). Physiology and pathophysiology of heme: implications for kidney disease. J Am Soc Nephrol *18*, 414-420.

Weitz, S.H., Gong, M., Barr, I., Weiss, S., and Guo, F. (2014). Processing of microRNA primary transcripts requires heme in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 1861-1866.

Ye, W., and Zhang, L. (2004). Heme deficiency causes apoptosis but does not increase ROS generation in HeLa cells. Biochem Biophys Res Commun *319*, 1065-1071.

Zhang, L., ed. (2011). Heme Biology: The secret life of heme in regulating diverse biological processes. (Singapore: World Scientific).

Chapter 4: Beyond the Canonical Hairpin: A structural motif in a common shRNA backbone

In searching for a minimal pri-miRNA that still processed effectively, Michael Faller discovered trimming off sequences around the canonical pri-miR-30a hairpin cause processing defects in vitro. Since pri-miR-30a is used as a common shRNA backbone, I wanted to test if this surrounding sequence had functional implications on shRNA processing and knockdown efficiency. I directed and mentored Timothy Huang, a visiting Amgen Scholar and Ayesha Hamid, a summer student from Santa Monica College, and both worked hard to contribute data to this project. We discovered the presence of a flanking helix which enhanced knockdown efficiency of shRNA. Dr. Princess Gilbert, advised me on constructing a basic phylogenetic tree, and Javona Whitebear, a computer programmer from University of Michigan, performed the bioinformatics analysis on pri-miR-30a.

Introduction

Better understanding of the structural requirements for miRNA production in cells empowered researchers to design miRNAs mimics to knock-down proteins of interest. Short-hairpin RNA (shRNA) are a method of introducing small RNA in cells from DNA vectors for RNA interference (RNAi). shRNAs are widely used as tools for gene repression, and are being explored in clinical trials for therapeutic potential (1,2). shRNA can be "programmed" by replacing the stem of the hairpin with sequence targeting a new gene of interest. Silencing RNA (siRNA) are another commonly used RNAi based method to achieve gene knockdown. Each approach has costs and benefits. siRNA can induce very potent knockdown effects, however they have been shown to saturate RNAi protein machinery causing toxicity and potentially unintended signal through the immune response pathway (3,4). Furthermore, siRNA can be diluted over time as cells divide thereby, their knockdown effect diminishes within a few days. shRNA can be integrated into the genome and therefore the gene knockdown is more stable

over time. Further, the toxicity problems associated with siRNA are not observed with shRNA as the endogenous proteins control the expression and processing. In 2005, Zhou et al. showed that using wild-type pri-miRNA sequence as a shRNA vector was significantly more effective than using a pre-miRNA hairpin based (first generation) design (5,6). Additionally, it was found that using Pol II transcript vs Pol III was more reliable in producing efficient knock-downs (7). Pri-miRNA based shRNAs (second generation) are referred to hereafter as shRNA^{miR}.

It is well established that the miRNA hairpin is evolutionarily conserved (8,9), however little research has been done on the potential functional roles of the structured sequence that surround the miRNA hairpin. Most shRNA backbones currently in use are based on the pri-miR-30a sequence. The pri-miR-30a backbone sequence has been progressively optimized for increased knock-down efficiency by several research groups. Systematic mutation and processing activity measurements showed that moving the EcoRI cloning site to a non-conserved region, thereby restoring some wild-type features such as a CNNC motif (10), enhanced the knockdown capacity 30-fold (miR-E) (11). This motif was shown to enhance association of the Microprocessor accessory protein, SFR3 to promote more efficient processing. Other research has indicated a minimal length of the flanking regions of 30 nucleotides on either side of the hairpin (4,12). note, small changes in the bulges of the stem seem to have little functional effect on shRNA efficiency (12).

One of the remaining confounding factors in shRNA^{mir} design is the somewhat arbitrary determination of the length of a pri-miRNA transcript. Researchers have used constructs that contain 100 – 300 nucleotides to represent the pri-miRNA transcript. A decade of biochemical data has identified structural features that define canonical pri-miRNA hairpin. These are generally accepted to be: a terminal loop, apical junction, a double stranded stem ~33 bp in length, a basal junction and unstructured regions that immediately surround the hairpin

(9,13,14). Much of this information was gathered through biochemical studies. The secondary structure largely defines what is identified as a pri-miRNA hairpin, therefore the pri-miRNA folding under experimental conditions will likely be a key variable in processing efficiency. However, predicting secondary structure in living cells in the context of full pri-miRNA gene and RNA binding proteins is a substantial challenge, therefore determining a clear set of rules for optimal shRNA in cells has been more difficult to define. For example, in optimization of shRNA^{miRs} the many studies conclude that long stretch of surrounding sequence do little to enhance shRNA^{miR} activity despite biochemical data to the contrary. This work investigates the functional contribution of sequence outside the canonical hairpin.

Results

Identification of a processing competent minimal hairpin

We wanted to determine the minimal pri-miRNA hairpin that processed well and could be used for *in vitro* processing assays. Therefore, we truncated the base of the pri-miR-30a hairpin to generate pri-miR-30a 80-mer, a little bigger than pre-miR-30a (83 nt), a 3'-truncation (118 nt) and a 5'-truncation (136 nt). Minimal pri-miR-30a was defined as the 3' and 5' mutations together (104 nt) (Fig. 1A). These constructs were transcribed and uniformly labeled with ^{32}P–UTP *in vitro.* The resulting labeled RNA was incubated with purified recombinant Drosha and DGCR8, which make up the essential protein components of the Microprocessor Complex (MC) (Fig. 1B). We observed proper processing from 150-nt pri-mir-30a termed "L1", however, processing activity was almost completely lost as the hairpin was truncated to 104 and 83 nt (Fig. 1B). In addition, truncations from the 3' or 5' side also resulted in a decrease in processing efficiency.

Disruption of the flanking helix impairs pri-miRNA processing *in vitro*

We observed that the 3' and 5' truncations disrupted a predicted helix directly flanking the pri-miR-30a hairpin and wondered if the pairing interactions might be important for proper processing. To assess the functional relevance of the flanking helix (f-helix), our group designed a series of RNA constructs based on pri-miR-30a. The 5' mutant disrupts the pairing interactions in the f-helix by mutating the 5' nucleotides that participate in binding. The 3' mutant disrupts the pairing from the 3' side. The compensatory mutant (C. mut) contains both the 5' and 3' mutations thereby restoring the pairing interactions, but with a different sequence from the wild-type (Fig. 2). These constructs were tested in the *in vitro* processing assay. The results showed clearly that disruption of the f-Helix from either the 5' or the 3' side, impaired MC processing and generated an alternative (shorter) pre-miRNA product (Fig. 2). Strikingly, restoration of the f-Helix restored proper pri-miRNA processing (Fig 2, lanes 12-13).

Disruption of the f-helix perturbs pri-miRNA processing in human cells

To test the effect of the f-helix in the context of living cells, shRNA targeting luciferase with the pri-miR-30a backbone were clone into the live-cell reporter described previously (15). Briefly, the reporter contains a bidirectional promoter that simultaneously drives the expression of eYFP and mCherry. The shRNA^{mir} sequences (WT, 5' mut, 3' mut and C.mut) were cloned into the 3' UTR of the mCherry expression cassette, therefore, processing causes destabilization of *mCherry* mRNA. pri-miRNA processing is measured as a decrease in the eYFP/mCherry ratio. This series of shRNA constructs were transfected into Tet-on U2-OS and HeLa cell lines. Results in U2-OS showed slight decreases in processing efficiency when the f-helix was disrupted from either the 5' or 3' side in both cell lines (Fig. 3A). The processing efficiency in HeLa cells did not seem to follow this same trend. The WT and the C. mut processed less efficiently than the 5' and the 3'mutants (Fig. 3B). This could be due to differences in expression of biogenesis proteins between cells lines.

The f-helix plays are role in shRNA^{mir} knockdown efficiency

To assess gene knockdown efficiency, Tet-on U2-OS cells that constitutively express firefly luciferase were transfected with the shRNA constructs. Cells were harvested and counted, and the same number of cells were lysed and assayed for luciferase signal. Disrupting the f-helix from with the 5' or 3' side resulted in large losses of luciferase knockdown. However, restoration of the pairing interaction in the f-helix restored shRNA knockdown efficiency (Fig. 3 C-D). This indicates that the f-helix is playing a functional role in knockdown efficiency in the U2-OS cellular context. In HeLa cells, opening the f-helix seems to promote processing which suggests that in some contexts the pairing interactions play in inhibitory role in processing

The f-helix shows evidence of conservations across species

Alignment of orthologous pri-miR-30a sequences using the human pri-miRNA as a reference showed evidence of evolutionary conservation across the canonical hairpin with high sequence identity throughout the evolutionary tree. Although not as highly conserved, a higher trend of identity ($100\% \geq \text{identity} \geq 30\%$) was observed for the region corresponding to a non-canonical helix bracketing the pri-miRNA hairpin. We used Genomic Evolutionary Rate Profiling (GERP++) analysis to examine the rate of rejected substitutions at each position for the alignment (16–18). GERP++ calculates a neutral substitution rate, and the number of observed substitutions for each nucleotide. The rejected substitution score (RS score) is calculated as the neutral substitution rate minus the observed substitution for a given alignment (17). GERP analysis revealed some evidence that nucleotide identity in this helical region was under selective pressure (Fig. 4A). GERP analysis can also give information about constrained structural elements, by clustering high scoring regions and testing the probability that these segments score higher significantly or merely by chance. Results from the constrained element search only returned the canonical pre-miRNA hairpin (Fig. 4B).

Discussion

While attempting to define the minimal pri-miRNA hairpin, we discovered that deletions of sequence surrounding the canonical pri-miRNA hairpin greatly reduced processing efficiency, prompting us to question a possible functional role for this sequence. We observed these sequences formed a predicted helix and wondered if this secondary structure played a role in processing and knockdown efficiency.

Disrupting the f-helix resulted in lower processing efficiency both *in vitro* and in U2-OS cells. Further, a disrupted f-helix caused a decrease in knockdown efficiency of luciferase. This suggests that pairing interactions directly flanking the canonical pri-miRNA hairpin alter processing efficiency. In addition, although the changes in shRNA processing efficiency where subtle, the resulting changes in knockdown efficiency were much more pronounced. These results seem to suggest that small changes in processing efficiency can translate into larger physiological changes.

After the human genome was sequenced in early 2003, the ENCODE projects preliminary report on 1% of the genome (30Mb) delivered a wealth of new insights into the genome. Contrary to the thinking of the time, they demonstrated that the most of the genome is undergoing active transcription, and a large percentage of these transcripts do not code for protein (18). Another intriguing observation was that biological functionality of a transcribed product is not related to the degree of evolutionary constraint. In fact, in the first report in 2007, they estimate ~50% of non-coding functional elements are unconstrained across all mammals (encode paper again). The f-helix appears to have some evidence of conservations across species especially in higher mammals. However, according to GERP++ analysis, is not under evolutionary constraint as a structural motif. This does not rule out the possibility that it is an emerging motif, or, as this work demonstrates, a functional structural motif.

The alternative cleavage products observed in the *in vitro* assay highlights that a pri-miRNA may be processed efficiently, but processing itself doesn't guarantee functionality. This is exemplified in the fact that 3' and 5' mutations still undergo processing, but an alternative pre-miRNA is produced indicated by a difference in size (Fig. 2B). Speculatively, determining the sequence on this alternative cleavage may reveal a change in the mature miRNA sequence, which would help explain the change in knockdown efficiency.

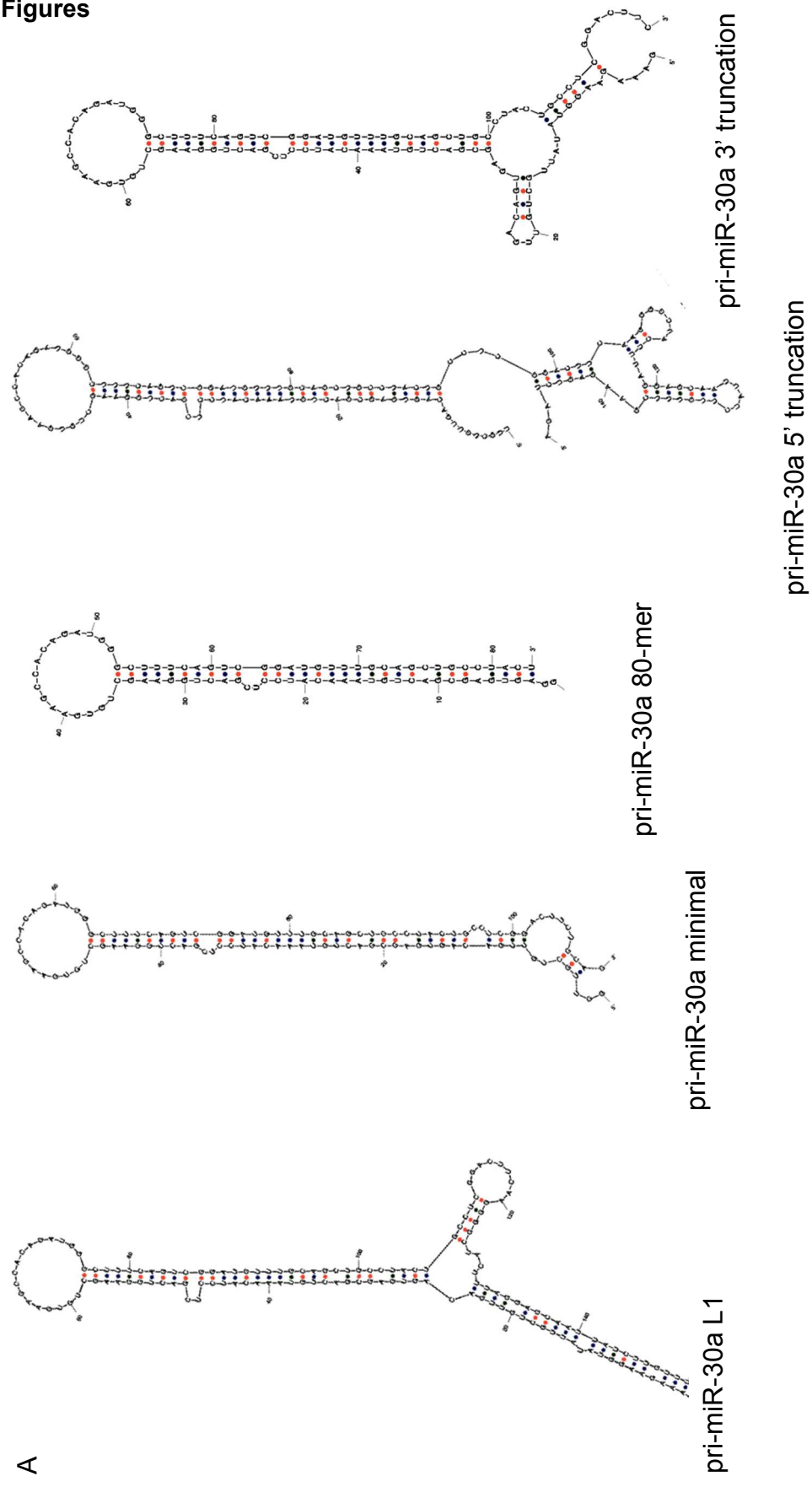
Our group has performed extensive secondary structure prediction of pri-miRNA hairpins ~450 nucleotides in length that indicate one cannot unambiguously identify the canonical pri-miRNA hairpin when longer stretches of sequence are included (data not shown). Recently, the Weinberg group published a similar observation, and demonstrated that a pri-miRNA that forms with sub-optimal (non-canonical) hairpins can still process and perhaps use several sequence motifs (10,19) to compensate for deviations from the canonical hairpin. It is possible that structural motifs could also serve the purpose of "normalizing" irregular pri-miRNA hairpins to facilitate their processing.

This study is not yet finished. Knockdown efficiency in HeLa is not addressed, nor is the differences in processing efficiency in different cell lines. Using a more physiologically "normal" context would be an important experiment to add. Additionally, the question of whether this enhancer will work with other hairpins should be answered. Constructs targeting different genes should be tested to see how broadly applicable the effects of the f-helix are.

In summary, we have identified a helix that seems to affect the processing efficiency of the well-studied pri-miR-30a. This structure is in the non-conserved region of pri-miRNA, much like the recent report of METTL3 sites (GGAC) that enhance processing efficiency by recruiting DGCR8 (20,21). This brings forward the interesting possibility that other structural enhancers

and/or regulatory elements could be in the sequence surround the pri-miRNA hairpin and could be exploited to improve shRNA scaffold design.

Figures


A
pri-miR-30a L1
pri-miR-30a minimal
pri-miR-30a 80-mer
pri-miR-30a 5' truncation
pri-miR-30a 3' truncation

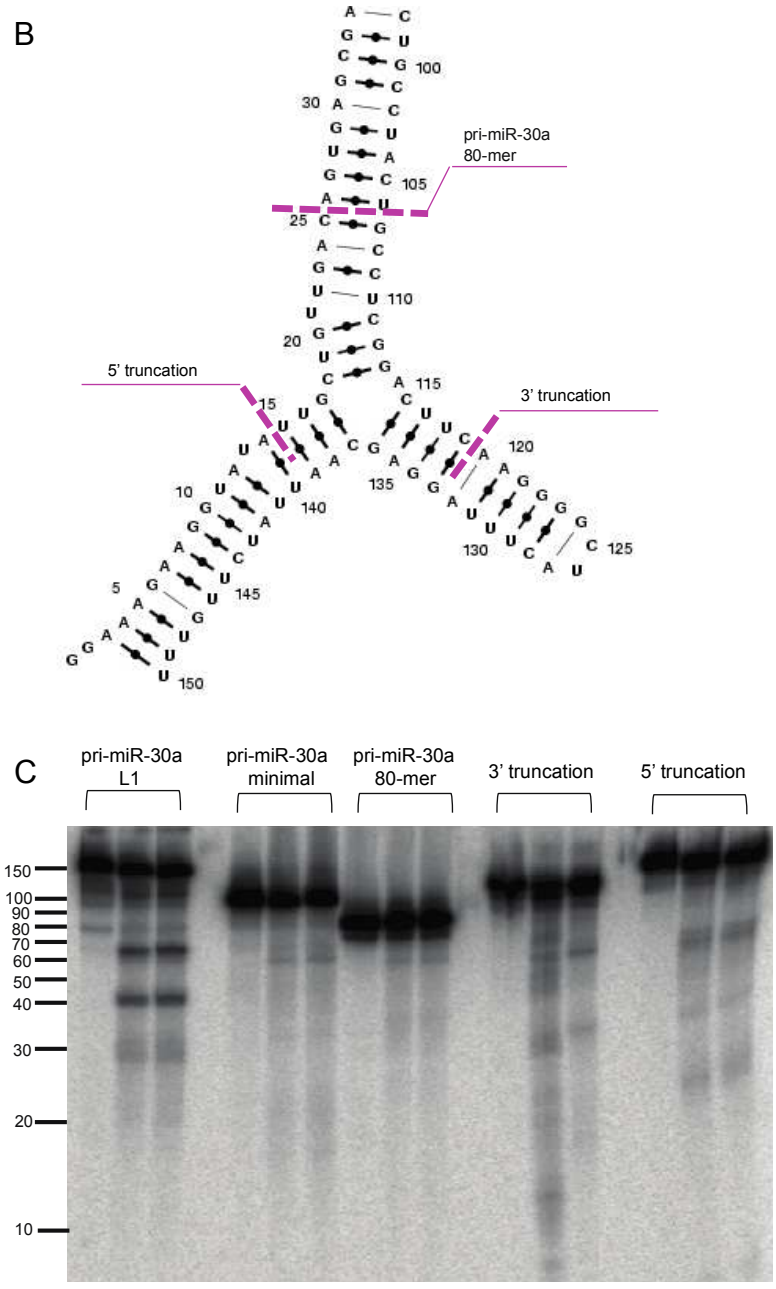


Figure 1. Identification of a pairing interaction in pri-miR-30a important for processing.

A) Secondary structure prediction of constructs used in the live-cell assay. B) Secondary structure prediction with emphasis on the basal junction and flanking regions. The truncation positions are

indicated by pink dashed lines. C) *In vitro* processing assay showing processing activity for pri-miR-30a and poor processing for the other truncations.

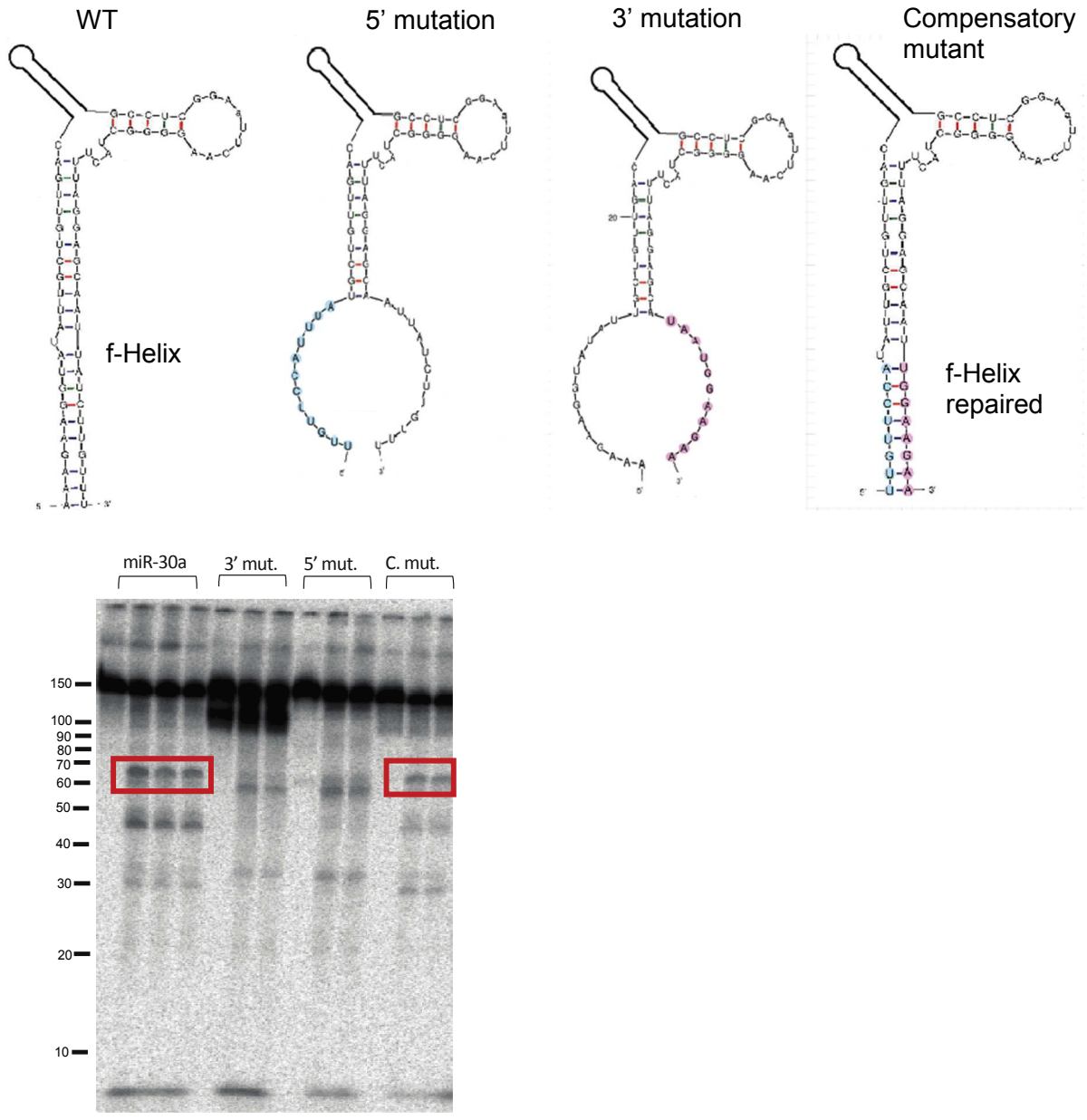


Figure 2. Disruption of the f-helix perturbs disrupts proper pri-miRNA processing. A) Secondary structure predictions of the f-helix disruption constructs. The 5' and the 3' mutations are highlighted in blue and pink respectively. B) Each set of lanes contain RNA alone, RNA+MC with 50 and 100 nM DGCR8 concentration. Red boxes indicate proper processing of miR-30a.

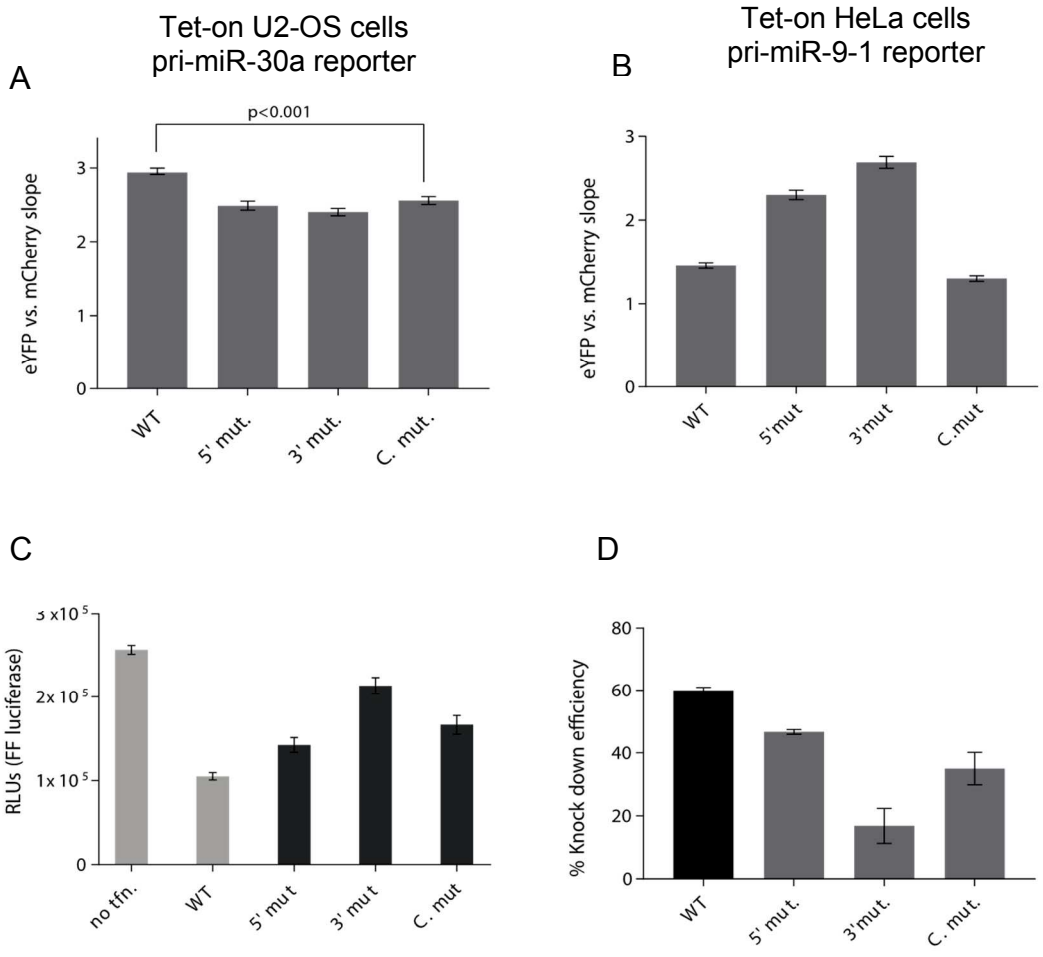


Figure 3. The f-helix plays a role in shRNA processing and knockdown efficiency.

A) In U2-OS cell the f-helix plays an enhancing role in processing efficiency while in HeLa cells (B) the paired region inhibits processing efficiency. C-D) A dual-luciferase assay shows that the f-helix aids in effective luciferase knockdown.

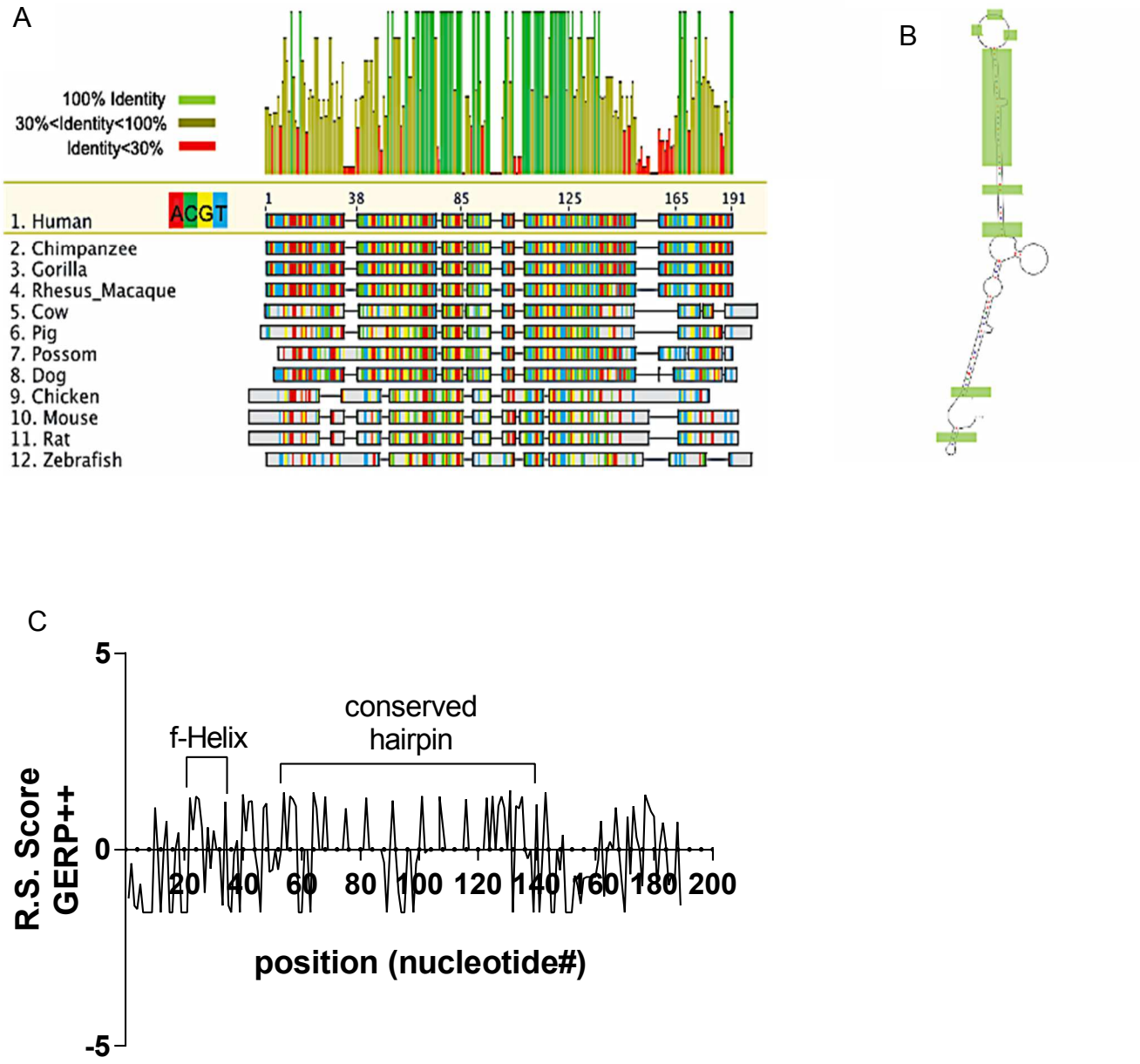


Figure 4. The f-helix shows some conservation across species. A) pri-miR-30a sequences included 60 nucleotides beyond the annotated miRbase hairpin were aligned showing some sequence identity in the f-helix region. B) GERP++ analysis of the alignment showed there is a positive rate substitution score in the f-helix.

Materials and Methods

Protein expression and purification

Briefly, DGCR8 proteins were expressed in *E. coli* strain BL21(DE3) CodonPlus. Cells were grown to late-log phase. Expression was induced with 1 mM IPTG for 4 hr at 37°C. δ-Aminolevulinic acid was added to a final concentration of 1 mM to facilitate heme loading to the protein. Cells were harvested by centrifugation at 5000xg, for 5 min at 4°C. Protein was initially purified using a SP cation-exchange column (GE Healthcare). Column equilibration and loading were performed using Buffer *A*, which contained 100 mM NaCl and 20 mM Tris pH 8.0. Linear gradient elution was performed with Buffer *B* containing 2 M NaCl and 20 mM Tris pH 8.0. Fresh dithiothreitol (DTT) was added to both buffers at a final concentration of 1 mM. Fractions containing the DGCR8 protein were concentrated in a centrifugal concentrator with MWCO of 30 kDa (Millipore), and loaded onto a Superdex-200 size exclusion chromatography column (GE Healthcare) equilibrated with a SEC buffer containing 400 mM NaCl and 20 mM Tris pH 8.0. Peak fractions were again concentrated and analyzed using a SDS 12%-polyacrylamide gel to determine purity.

Cell culture and Live-cell pri-miRNA processing reporter assay

Tet-on Hela and U2-OS cells were grown at 37°C with 5% CO_2 in DMEM medium containing 4.5 g/L glucose, L-glutamine and 110 mg/L sodium pyruvate supplemented with 10% Tet-System-Approved FBS (Clonetech Laboratories). For imaging, cells were first passaged into 3.5-cm dishes with a glass coverslip bottom. Cells were transfected at ~80% confluency with Effectene (Qiagen), and the expression of reporters was induced by addition of 1 μg/mL doxycycline. After ~20 hr, cells were

washed and exchanged into DMEM without phenol red. Fluorescence imaging was performed using a Nikon Eclipse Ti microscope with an EMCCD camera (Andor iXon). We used a 10x objection lens and an exposure time of 0.3 ms with gain at 0. Fluorescence was detected with filter cubes eYFP (510 ± 10 nm band pass excitation and 535 ± 15 nm band pass emission) and mCherry (535 ± 25 nm band pass excitation and 610 ± 10 nm band pass emission). Images were exported as 16-bit tiff files. Images were analyzed using mostly as described (19), except that a new Matlab (The MathWorks) program called "Jose2Dcellcounter" was written for integrating fluorescence intensities of individual cells in a streamlined fashion. Using the image analysis toolbox in Matlab, our program first performs segmentation to determine the locations of individual nuclei from pairs of eYFP and mCherry images acquired from the same field of view. A mask is created such that pixels with intensity above a user-defined threshold are considered to be in nuclei and are assigned a value of 1. The values of all other pixels are set to 0. We choose threshold values that maximize the number of cells detected. The perimeter of a nucleus is determined for each image of a pair separately, and then a common mask is chosen for the two images based on the following criteria. If the same cell is found in both images, the mask with smaller perimeter is chosen to avoid including cytosols, which sometimes have low fluorescent signals. If a cell is detected in only one image, the perimeter is kept, allowing for very dim cells that are not detected in one channel to be included. The center of the final perimeter for each cell is determined. After the nuclear perimeters are determined, downstream analyses are performed on the original images. A background fluorescence value is calculated for each image by averaging the intensities of the pixels outside the identified nuclei. The background is then subtracted from the intensities of all pixels.

The total intensity of an individual cell is calculated as the sum of the pixel intensities within the perimeter. Dead cells can be removed manually. A list of total eYFP and mCherry intensities for individual cells is exported to Microsoft Excel, scatter plots are generated and linear fit of eYFP versus mCherry signals gives the fluorescence slopes.

Dual-Luciferase Assay

Tet-on U2-OS_luc cells were transfected with the shRNA constructs. At 20 hr posttransfection cells were harvested using 0.05% trypsin-EDTA and counted with a Countess II automated cell counter. Promega Dual-Luciferase Assay kit E19110 and a Monolight 2010 luminometer were used to measure luciferase signals. We standardly incubate cells under the passive lysis condition (list the buffer composition if known) for 20 min and measure luciferase signal for 5 seconds,

GERP analysis:

GERP++ was downloaded from:

http://mendel.standfor.edu/SidowLab/downloads/gerp/index.html

Sequences were downloaded from UCSC genome browser and trimmed to 191 nucleotides for each pri-miRNA. Sequence alignment was constructed in Geneious R9 using the Clustal method. Phylogenetic tree was constructed from the alignment using the UPGMA algorithm.

Author Contributions

Michael Faller performed the *in vitro* truncation analysis of pri-miR-30a. Jen Quick-Cleveland, Timothy Huang and Ayesha Hamid performed all the cellular experiments and luciferase knock-down assays. Jen Quick-Cleveland generated the alignment and phylogenetic tree for GERP++ and Javona Whitebear did the GERP++ computational analysis. Michael Faller, Feng Guo, Jen

Quick-Cleveland designed the experiments and Jen Quick-Cleveland and Feng Guo wrote the manuscript.

Acknowledgements

This work was supported by NIH grant GM080563 to F.G., Ruth L. Kirschstein National Research Service Award GM007185 to J.Q., T.H. was supported by the Summer Program for Undergraduate Research (SPUR) and A.H. was supported by the UCLA/SMC Science Research Initiative. We would like to thank Dr. Princess Gilbert for her discussions about generating the alignment for GERP++ analysis.

References

1. Senzer N, Barve M, Kuhn J, Melnyk A, Beitsch P, Lazar M, et al. Phase I trial of "bi-shRNAi(furin)/GMCSF DNA/autologous tumor cell" vaccine FANG in advanced cancer. Mol Ther. Nature Publishing Group; 2012;20(3):679–86.

2. Burnett JC, Rossi JJ, Tiemann K. Current progress of siRNA/shRNA therapeutics in clinical trials. Biotechnol J. 2011;6(9):1130–46.

3. Moore CB, Guthrie EH, Huang MT, Taxman DJ. RNA Therapeutics. 2010;629(2):1–15.

4. Bofill-De Ros X, Gu S. Guidelines for the optimal design of miRNA-based shRNAs. Methods [Internet]. 2016;103:157–66.

5. Zhou H, Xia XG, Xu Z. An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi. Nucleic Acids Res. 2005;33(6):1–8.

6. Hinton H.M., Wise T.G., Cottee P.A.,Doran T.J. Native microRNA loop sequences can improve short hairpin RNA processing for virus gene silencing in animal cells. J RNAi Gene Silenc. 2008:4(1);295-301.

7. Xia XG, Zhou H, Samper E, Melov S, Xu Z. Pol II-expressed shRNA knocks down Sod2 gene expression and causes phenotypes of the gene knockout in mice. PLoS Genet. 2006;2(1):73–80.

8. Berezikov E, Guryev V, Van De Belt J, Wienholds E, Plasterk RHA, Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. Cell. 2005;120(1):21–4.

9. Yeom K-H, Lee Y, Han J, Suh MR, Kim VN. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Res. 2006 Jan;34(16):4622–9.

10. Auyeung VC, Ulitsky I, McGeary SE, Bartel DP. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell . 2013 ;152(4):844–58.

11. Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, et al. An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. 2013;5(6):1704–13

12. van den Berg FT, Rossi JJ, Arbuthnot P, Weinberg MS. Design of Effective Primary MicroRNA Mimics With Different Basal Stem Conformations. Mol Ther Acids. IOP Publishing; 2016;5(1).

13. Zeng Y, Yi R, Cullen BR. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. 2005;24(1):138–48.

14. Zeng Y, Cullen BR. Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences. 2005;280(30):27595–603.

15. Weitz SH, Gong M, Barr I, Weiss S, Guo F. Processing of microRNA primary transcripts

requires heme in mammalian cells. Proc Natl Acad Sci U S A. 2014;111(5):1861–6.

16. Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 2005;15(7):901–13.

17. Davydov E V., Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010;6(12).

18. Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung W-K, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007;447(7146):799–816.

19. van den Berg FT, Rossi JJ, Arbuthnot P, Weinberg MS. Design of Effective Primary MicroRNA Mimics With Different Basal Stem Conformations. Cit Mol Ther Acids. 2016;5.
20. Lee H, Goodarzi H, Halberg N, Tavazoie SF. HHS Public Access. 2015;519(7544):482–5.

21. Knuckles P, Carl SH, Musheev M, Niehrs C, Wenger A, Bühler M. RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat Struct Mol Biol. Nature; 2017;(June).

Chapter 5: Cancer-associated mutations render the DGCR8 protein defective in pri-miRNA processing.

The ideas that fueled this project came from a Guo lab idea meeting where it was proposed we investigate the Catalogue of Somatic Mutations in Cancer for any interesting DGCR8 mutants found in tumors. Grant Shoffner and I selected mutations to test, and designed the experiments in this chapter. We both directed and mentored undergraduates Kristina Solorio, Kelsey O'Donahugh and Caitlyn Pickard who all contributed to the project. Together with my students, I performed all the cellular experiments which demonstrated some of the DGCR8 mutants found in tumors were processing-defecting. We also cloned and characterized the E518K mutant, a mutation found in ~3% of Wilms tumors. I found that two of the mutants located in the Rhed had decreased pri-miRNA and heme-binding affinities. Under my direct mentoship, Kristina Solorio and Kelsey O'Donahugh were able to produced some of the data used in the final manuscript. Our contribution linked some of the DGCR8 mutants found in cancer directly to pri-miRNA biogenesis defects, thereby indicating they could be potential drivers of tumorigenesis. Grant and myself wrote the manuscript with assistance from Dr. Guo, and it will be submitted for publication.

Introduction

The genesis and development of cancer remains an engrossing biological problem. The intricacy of the molecular pathways involved has made progress in research and successful treatment challenging. Data collected on cancer tissues has led to the insight that there are many mutations in cancerous cells (1–3). Untangling disease-driving versus spurious mutations presents a daunting challenge. This is further complicated by that fact that collective mutations often differ between patients, even with the same kind of cancer (4,5). To overcome the challenge in treating heterogenous tumors, it is important to identify potential driver mutations

for pathogenesis. Estimates for mutations that drive cancer development range from 10,000 (5) to, more recently, as little as 3 (6), and vary depending on cancer type.

miRNA dysregulation is a hallmark of malignant cells (7–9). Coukos *et. al* discovered that there was a high frequency of genomic alterations at miRNA loci (10). About 50% of miRNA genes were found in cancer-associated genomic regions linking miRNAs with cancer pathogenesis (11). Specific miRNAs are often overexpressed or suppressed in cancer cells. Additionally, there is a global down-regulation of miRNA expression in many cancers (7,8). Some miRNAs, such as let-7, miR-143/145, miR-200 are tumor supressors (11–13). Others, such as the miRNA cluster miR-17~92 and miR-155 are oncogenic (14). In non-cancerous cells, global reduction of miRNA expression has an overall effect of favoring cellular transformation and tumorigenesis (7,15) .

miRNA biogenesis proteins undergo changes that are strongly implicated in cancer development. In the canonical miRNA maturation pathway, primary transcripts (pri-miRNAs) are cleaved in the nucleus by the Microprocessor Complex (MC), which contains the RNase Drosha and its obligate RNA-binding partner DGCR8. The resulting precursor miRNAs (pre-miRNA) leave the nucleus via Exportin-5. Once in the cytoplasm, the RNase Dicer performs another cleavage step. The single-stranded ~22-nt mature miRNAs are incorporated into effector Argonaute proteins forming the *R*NA- *i*nduced *s*ilencing *c*omplex (RISC.). RISC uses the mature miRNA as a guide to bind cognate mRNA ultimately leading to repression of the target (16–18). Expression of Drosha and Dicer is often decreased in ovarian cancer specimens, and such decreases correlate with poor prognosis (19). Reduction of DGCR8 or Dicer expression enhances cellular transformation and tumorigenesis in mice. This demonstrates a direct contribution of abnormal miRNA biogenesis to tumor formation (15,20). Hypoxia, another hallmark of cancerous cells, causes coordinated repression of miRNA

biogenesis proteins Dicer, Drosha, TARPB2, XPO5 and DGCR8 in breast cancer cell lines (21). Drosha and DGCR8 mutations and deletions appear in 15% of Wilms tumors, a cancer that originates in the kidney and is the most common pediatric renal malignancy in the United States (22,23). DGCR8 specifically plays a role in prostate cancer progression. In *PTEN* knock-out mouse model for prostate cancer loss of DGCR8 expression inhibits tumor progression (24). Expression of the miRNA biogenesis proteins varies depending on cancer type; in prostate cancer these proteins are upregulated, while in most lung and ovarian cancers they are downregulated (24). This tissue-specific regulation and role of miRNA in cancer cells makes dissecting general molecular mechanisms even more challenging.

DGCR8 is a unique RNA-binding protein. It contains two C-terminal double-stranded RNA-binding domains (dsRBDs), and an RNA-binding heme domain (Rhed) that is required for recognition and processing of pri-miRNA substrates (Fig. 1A) (25,26). Using the Rhed, DGCR8 forms a dimer to bind a ferric heme cofactor (27). DGCR8 binds Fe(III) heme with $>10^7$-fold higher affinity than that for the Fe(II) form and, importantly, only Fe(III) heme can activate pri-miRNA processing activity (28). DGCR8 ligates heme iron using two cysteine side chains, one (C352) contributed by each subunit. This heme ligation configuration generates an electron-rich environment for the heme iron that confers extraordinary redox specificity to DGCR8 (28). In addition, DGCR8 is regulated by phosphorylation(29), sumoylation (30), acetylation (31), caspase-mediated proteolytic cleavage (32), feedback cleavage of pri-miRNA-like hairpins in its mRNA (33,34) and competition from non-pri-miRNAs (35). Therefore, the DGCR8 proteins is a point of convergence between many regulatory pathways that control miRNA biogenesis.

The motivation of the current study are two-fold. First, since DGCR8 is frequently mutated in tumors, we wanted to investigate whether these changes in DGCR8 could cause pri-miRNA processing defects. Second, biochemical characterization of cancer-associated DGCR8

mutations would shed light on the structure-and-function relationship of this protein. Using a combination of cellular and biochemical assays, we show that several cancer-associated point mutations alter DGCR8 properties and thereby disrupt pri-miRNA processing. Some of the mutations highlight the importance of the Rhed in DGCR8 function.

Results

Tumor-derived point-mutations in DGCR8 cause pri-miRNA processing defects

To investigate the susceptibility of DGCR8 to genetic inactivation in tumors, we searched the Catalogue of Somatic Mutations in Cancer (COSMIC) for mutations in the DGCR8 locus. From the 138 mutations annotated in the database, we observed that the majority (70%) arose from carcinomas, with cancers of intestinal origin occurring at the highest frequency (22%) (Fig. 1B). When classified by their effects on protein coding,11% resulted in frameshift or nonsense mutations (Fig. 1B, right panel), clearly deleterious to protein function. We noticed that 34% of the mutations produced missense substitutions. It is not clear whether these mutations affect DGCR8 function and contribute to oncogenesis.

We elected to analyze seven missense DGCR8 mutations for their functional implications (Fig. 1A). Three mutations are located in Rhed, including K289E (change in charge), G336E (change in flexibility and charge), and F448L (change in aromaticity). Four mutations in dsRBD1 (E518K, R527C, R527H and R570Q). E518K is a hot-spot mutation that occurs in about 3% of Wilms Tumors (22,23). Although E518K has been shown to alter miRNA expression profile, the underlying mechanism is unknown.

Comparing protein sequences across multiple homologs revealed that all sites of mutation are highly conserved, with the exception of R527 (Fig. 1C). For those mutations falling within crystal structures of DGCR8, we determined their spatial proximity to residues with known functional significance (Fig. 1D). G336 co-localizes with R322 and R325 on the surface of the

DSD (Fig. 1D, left panel) (36,37). We have previously shown that these basic residues are involved in binding pri-miRNA (25). Interestingly, E518, R527, and R570 on dsRBD1 are not spatially close to the previously reported RNA-binding patches on helix 2 (Fig. 1D, right panel, labeled as D1H2 and D2H2) (38).

We first tested the effects of tumor-derived DGCR8 mutations using a previously-developed live-cell pri-miRNA processing reporter assay (Fig. 2A) (39). The reporter uses a bi-directional promoter that drives the expression of two fluorescent proteins, eYFP and mCherry. We inserted a pri-miRNA hairpin into the 3'UTR of the mCherry expression cassette while eYFP serves as normalization. Cleavage of the pri-miRNA moiety by the MC results in a decrease in mCherry protein expression as the fusion mRNA is destabilized. Slopes from linear fits to the mCherry vs eYFP intensities of individual cells indicate pri-miRNA processing efficiency. Overexpression of wild-type DGCR8 causes a significant increase in processing efficiency, providing a convenient way to test mutants (Fig. 2B). To assure generality of our conclusions, we employed two pri-miRNA reporters (pri-miR-9-1 and pri-miR30a) in two human cell lines (HeLa and U2OS), respectively.

The reporter assays showed that four of the seven DGCR8 mutations cause severe defects in pri-miRNA processing in both HeLa and U2-OS Tet-on cell lines. In HeLa, overexpression of wild-type DGCR8 nearly doubles the mCherry vs eYFP slope (Fig. 2B). While R527C, R527H and R570Q are as active as the wild type, K289E and G336E are substantially less active. Expression of E518K and F448L did not increase pri-miRNA processing efficiency compared to the empty-vector control. F448L even causing a slightly decrease over the vector control, suggesting potential a dominant negative effect. In U2OS cells the same trend, though less pronounced, was observed. Overexpression of wild-type DGCR8 results in a 30% increase in mCherry vs eYFP slope and, K289E, G336E, F448L and E518K are severely defective in pri-

miRNA processing. The overexpression of F448L again drops the processing efficiency to below the endogenous level. Immunoblot analysis of transfected cells showed that the DGCR8 mutants were expressed at levels similar to the wild type (Fig. 2B), ruling out the possibility that the reduced pri-miRNA processing efficiency was due to uneven protein expression levels.

E518K disrupts the folding of DGCR8

To investigate why the cancer-associated mutations are defective in pri-miRNA processing, we engineered these mutations into a bacterial expression plasmid for a truncated DGCR8 construct called NC1, which contains all domains important for biochemical pri-miRNA processing activity (Fig. 1A). Using a previously established protocol, we expressed and purified these mutants, except for E518K. This mutant is expressed as an insoluble form under our standard condition. We screened number of salt concentrations and adjuvants but the protein remained largely insoluble. Therefore, we conclude that E518K disrupts the folding of DGCR8. This is consistent with crystal structure of DGCR8 core, which shows that E518 is involved in an extensive hydrogen-bonding network.

F448L abolishes DGCR8 binding to the essential heme cofactor

In contrast to purified NC1 wild-type and the K289E and G336E mutants, F448L mutant did not display the yellow-brown color characteristic of the Fe(III) heme cofactor. The wild-type NC1 protein has intense absorption peaks arise from the ligation of Fe(III) heme by the two cysteine side chains (28). However, the electronic absorption spectrum of purified NC1 F448L did not have the signature heme bands at 366, 450 and 556 nm (Fig. 3B). Furthermore, we attempted to reconstituted the NC1 F448L-heme complex by incubating the apo form of the protein with Fe(III) heme, but failed to observe any absorption peaks corresponding to specific DGCR8-heme interaction. Therefore, we conclude the F448L mutation abolishes heme binding to DGCR8 (Fig. 3B).

Next, we tested the activity of the purified NC1 mutants using pri-miRNA processing assays reconstituted with purified recombinant Drosha and ^{32}P-labeled substrates (Fig. 3C). Consistent with the results from cellular assays, we observed greatly attenuated activity for F448L, with very little pre-miR-9-1 and pre-miR-21 produced. Substantial amount of pre-miR-30a was generated by this mutant but still much less that that from the wild type. This is consistent with pri-miR-30a being less sensitive to DGCR8 defects in *in vitro* activity assays as previously reported (25). In contrast, NC1 G336E is only slightly less active than the wild type and NC1 K289E is just as active as the wild type. These observations are consistent with our previous report that the reconstituted pri-miRNA processing assay is much more forgiving than the cellular assay (25), probably due to the absence of other RNAs as competitors and other nuclear proteins as competitors. Overall, the biochemical assays confirm that F448L mutant has severe pri-miRNA processing defects. This is likely to be caused by by the disruption in heme binding, which would also alter MC complex formation as previously reported (25).

F448L also disrupts DGCR8's interaction with Drosha

To determine if the cancer-associated mutations might decrease DGCR8 activity by interfering with MC assembly, we developed a pull-down assay. This assay examines the interaction between purified recombinant N-FLAG-NC1 (N-flag-NC1) and His_6-Drosha$^{390-1374}$. Drosha is selective precipitated by N-flag-NC1 (compare lanes 11 and 5 in Fig. 4A). Their co-precipitation does not depend on pri-miRNA (compare lane 11 with lanes 14 and 8), indicative of a direct interaction. This observation is consistent with previous biochemical analyses (40,41) supporting a direct interaction between DGCR8 and Drosha. We then tested the DGCR8 mutants using the pull-down assay. We observed that while the N-flag-NC1 K289E and G336E mutants associate with Drosha to the same extent as the wild type, F448L fails to do so (Fig. 4B). Disruption of DGCR8-Drosha interaction, and interferance with the DGCR8-heme

interaction provides clear rationales for the inactivity of F448L mutant. Biochemical assays and crystal structure indicates that Drosha directly binds DGCR8 in the CTT region (41,42) Our results suggest that heme-binding to DGCR8 may be required for strong association with Drosha. It is possible that the Rhed provides an additional binding sites for Drosha.

K289E and G363E reduce affinities for heme and pri-miRNA and destabilize the structure

We used quantitative assays to measure how the K289E and G336E mutations may affect DGCR8's affinities for heme and pri-miRNAs, as well as the protein's structural stability. To examine changes in heme-binding affinity, we performed a kinetic heme-dissociation assay in which purified Rhed mutants were incubated with a large molar excess of a heme scavenger protein, apomyoglobin. Heme molecules that dissociate from DGCR8 are quickly picked up by apomyglobin, which has a K_d of 10^{-15} M for heme, and becomes unavailable to bind back to DGCR8. The heme transfer results in a decrease in the absorbance at 450 nm (A_{450}), and a concurrent increase in absorbance at 409 nm (A_{409}) indicating formation of ferric myoglobin.

We previously showed that DGCR8 binds Fe(III) heme extremely tightly, with no visible changes in A_{450} observed for human NC1 over 4-5 days (28). Here we measured changes in heme dissociation rate in the context of the Rhed. In almost 3 days, A_{450} had only decreased by very little (Fig. 5A). In contrast, both K289E and G336E Rhed mutants lose heme much faster. By fitting A_{450} over time with a one-phase exponential decay equation, we deduced the half-lives ($t_{1/2}$) and dissociation rate (k_{off}) to be 16.3 hrs and 1.70 x10^{-5} s^{-1} for K289E, and 16.1 hrs and 1.73 x10^{-5} s^{-1} for G336E. To measure mutational effects on pri-miRNA-binding affinity, we used a filter binding assay to measure the dissociation constant (K_d) for each mutant in the context of Rhed. Three pri-miRNAs are used. As we have shown in the past, wild-type NC1 binds all three pri-miRNAs with K_d values 10-16 nM. The affinities of NC1-K289E for pri-miR-21 and pri-miR-30a are similar to the wild type, but the affinity for pri-miR-9-1 is decreased, with K_d of 36 nM

(Figure 6A). NC1-G336E binds pri-miR-30a with a wild-type-level affinity, but had lower affinity for both pri-miR-9-1 and pri-miR-21, with and average K_d value of 52 and 50 nM, respectively. Importantly, the K289E and G336E mutations render the Rhed unable to bind the pri-miRNA substrate without the help of the dsRBDs (Table 1).

We also investigated whether the K289E and G336E mutations reduce DGCR8 structural stability. We performed the thermo-melting assays to compare the mutants with wild-type in the context of NC1. In these assays, A_{450} is monitored to assess association with heme and thereby serves as an indicator of proper protein folding. We observed that the thermostability of both K289E and G336E mutants is decreased relative to the wild type (Fig. 6A). Although we were not able to determine their exact melting temperatures due to substantial background scattering from aggregated proteins, visual inspection of the melting curves indicates that K289E and G336E destabilize the NC1 proteins by 2-3˚C. The effects of G336E are more pronounced, as below 50˚C NC1 G336E loses A_{450} more quickly than the wild type and K289E.

K289E and G336E are located within the independently folded dimerization sub-domain (DSD) of the Rhed. The DSD expresses very well in *E. coli* and has a crystal structure available (22). We analyzed DSD mutants by size-exclusion chromatography (Fig. 6B). Whereas wild-type DSD elutes at 59.2 ml, we observed a decreased elution volume (~56.3 ml) for both K289E and G336E that was still greater than the void volume of the column. We believe that the apparent increase in elution volume is caused by the dimeric DSD mutants unfolding, effectively increasing the hydrodynamic radius of the molecule.

Taken together, our biochemical characterization of the K289E and G336E mutants revealed moderate reductions in heme binding, pri-miRNA binding and structural stability. Any

one or combination of these changes could result in the pri-miRNA processing deficiency observed in cells. The biochemical effects of G336E are more pronounced, consistent with this mutant being more defective than K289E.

Discussion

Identifying driver mutations that favor uncontrolled growth and prevent apoptosis is critical for understanding cancer development and for formulating novel treatment strategies. A great example concerns B-Raf, a serine/threonine kinase involved in cell-growth signaling pathways. A point mutation, V600E, constitutively activates B-Raf and occurs in ~50% of melanomas (43)(44,45). The drug trametanib has been developed to treat patients carrying this mutation (46). Because it has been demonstrated that disruption of the miRNA maturation pathway leads to a pro-growth phenotype and the development of tumors (2), we analyzed tumor-associated DGCR8 mutations. We found that, out of 138 mutations in the COSMIC database, 11 cause frameshifts that most likely result in loss of function. Four out of the seven point-mutations we selected for experimental testing, are severely defective in pri-miRNA processing. Thus, a substantial fraction of the DGCR8 tumor-associated mutations render the protein defective and are likely to directly contribute to pathogenesis.

These tumor-associated mutations are presumably somatic, and arise heterozygously during tumor development. They may cause pri-miRNA processing deficiency by the following mechanisms. First, losing 50% of functional DGCR8 proteins directly results in a relatively mild reduction in pri-miRNA processing efficiency, which has been shown to reduce the maturation of a small fraction of miRNAs, as observed in DiGeorge syndrome patients and *Dgcr8*$^{+/-}$ mouse model (19,47). However, such reduction in pri-miRNA processing has not been linked to tumor development. A second mechanism is that one mutated copy of DGCR8 can still dimerize with

the WT copy, which in turn will form a defective MC resulting in dominant negative effects. Consistent with such an expectation, expression of F448L in HeLa cells results in a pri-miRNA processing efficiency lower than the vector control. Third, these DGCR8 mutations could result in a very severe defect, even a complete loss in pri-miRNA processing when combined with another event that reduces or abolishes the expression of the second copy of *DGCR8* gene. Such changes have been observed in the Wilms tumor patients carrying the E518K mutation (22,23). Through these mechanisms, pre-cancerous cells have many paths to dysregulate miRNA maturation in favor of proliferation and tumorigenesis.

There has been a great interest to understand the mechanistic underpinnings responsible for the global repression of miRNA expression in cancer. The reported molecular explanations included high tendency of miRNA genes to be located in fragile genomic loci (48), loss of expression of proteins involved in miRNA biogenesis (19) (19,21), and abnormal signaling that regulates miRNA processing (49). We and others showed that DGCR8 mutations found in tumors can cause functional and structural defects that lead to miRNA processing deficits (22,23). Therefore, tumors multiple ways of achieving global miRNA repression. We recently identified a metalloporphyrin compound, cobalt (III) protoporphyrin IX, that can bind and hyper-activate the DGCR8 protein and compensate heterozygous deletion of the *DGCR8* gene (50). This compound offers a potential way of correcting the pri-miRNA processing defects in tumors carrying *DGCR8* mutations. We imagine that therapeutic strategies like this will become a common practice in the era of personalized medicine.

The DGCR8 mutations characterized here lend insight into how DGCR8 functions in pri-miRNA processing. Three of the defective mutants, K298E, G336E, and F448L are located in the Rhed, which our group previously demonstrated is required for processing activity in cells (39). We showed that this domain participates in binding pri-miRNA by binding conserved pri-

miRNA junctions (25). We also demonstrated that this domain contains a conserved caspase cleavage site and that caspase-mediated cleavage of Rhed inactivates DGCR8 (32). Still others have provided evidence that this domain is critical for pri-miRNA processing accuracy (26). Our work here highlights that the Rhed of DGCR8 appears to be particular sensitive to perturbation in tumors. F448L, K289E, and G336E weaken the heme-binding affinity to various degrees, reinforcing the physiological importance of DGCR8-heme interaction, though its exact function remains undefined.

Our data shows that both K289E and G336E differentially reduce the affinities for different pri-miRNA hairpins. Humans have ~2000 annotated pri-miRNA hairpins (51). The predicted secondary structures of human pri-miRNA are quite different from each other. Our observation suggests that DGCR8 interacts with individual pri-miRNA hairpins differentially. As such, we expect that the DGCR8 mutations affect some miRNAs more than others.

The heme-binding deficient F448L failed to associate with Drosha. This observation suggests that the Rhed and its association with heme are important for MC formation. The C-terminal tail of DGCR8 has been shown to mediate a low-affinity interaction with Drosha (26,42). However, the CTT-Drosha interaction appears to be insufficient to maintain the MC in our pull-down assays. Therefore, our study reveals a new function of the Rhed in mediating interaction between Drosha and DGCR8, in addition to binding pri-miRNAs (25,26) .

.

Figures

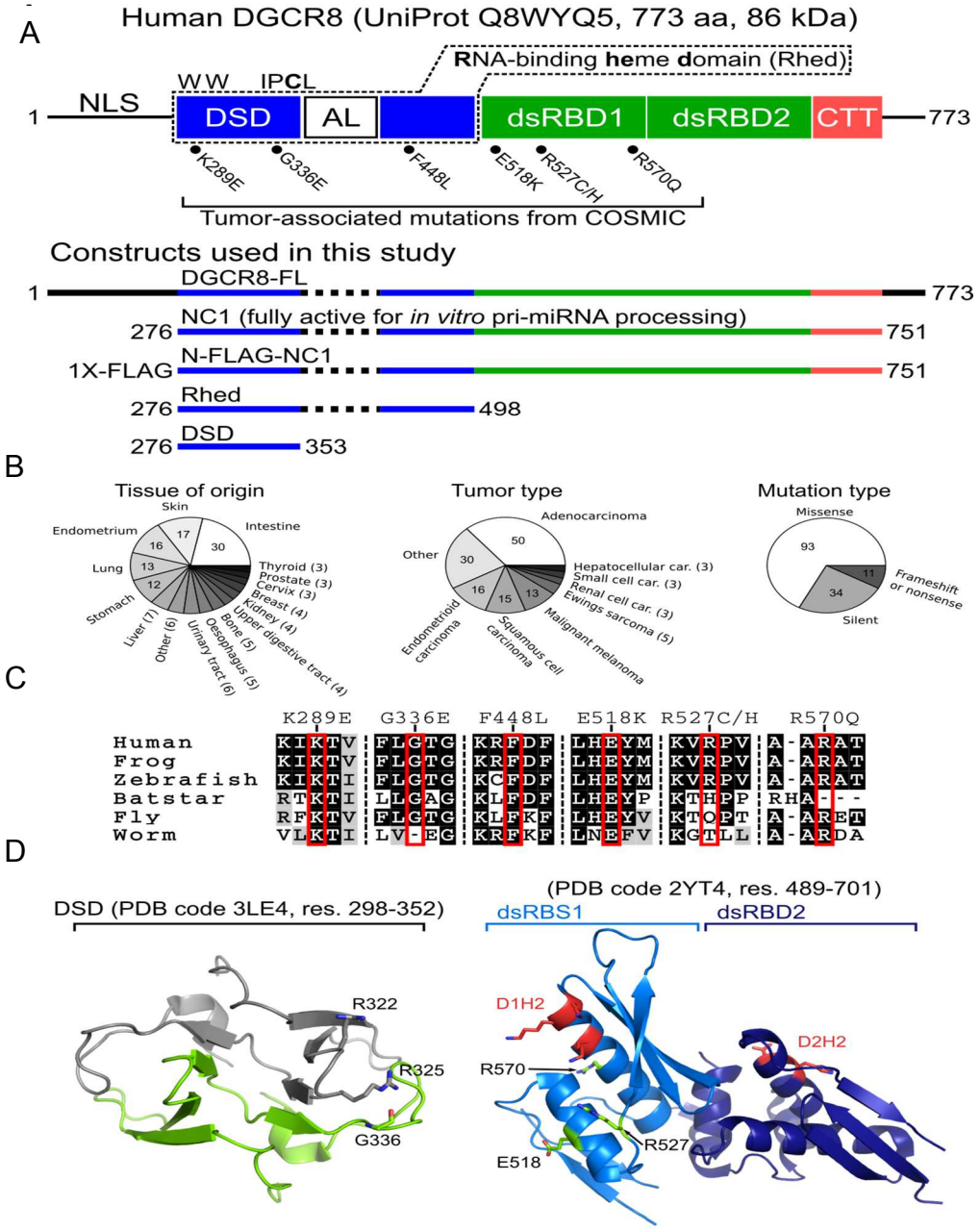


Figure 1. **Analysis of tumor-associated mutations in DGCR8**. (A) Domain structure of DGCR8 and expression constructs. The RNA-binding heme domain (Rhed) is shown as an N-terminal dimerization subdomain (DSD), a central acidic loop (AL), and a C-terminal subdomain. A WW motif and IPCL heme-binding site are labeled. (B) Classification of 138 somatic DGCR8 mutations observed in COSMIC database. (C) Multiple sequence alignment of DGCR8 homologs highlighting mutations considered in this

study. (D) Crystal structures of the DGCR8 DSD (left panel) and dsRBDs (right panel). Basic residues previously shown to interact with RNA are highlighted in red (D1H2 and D2H2)

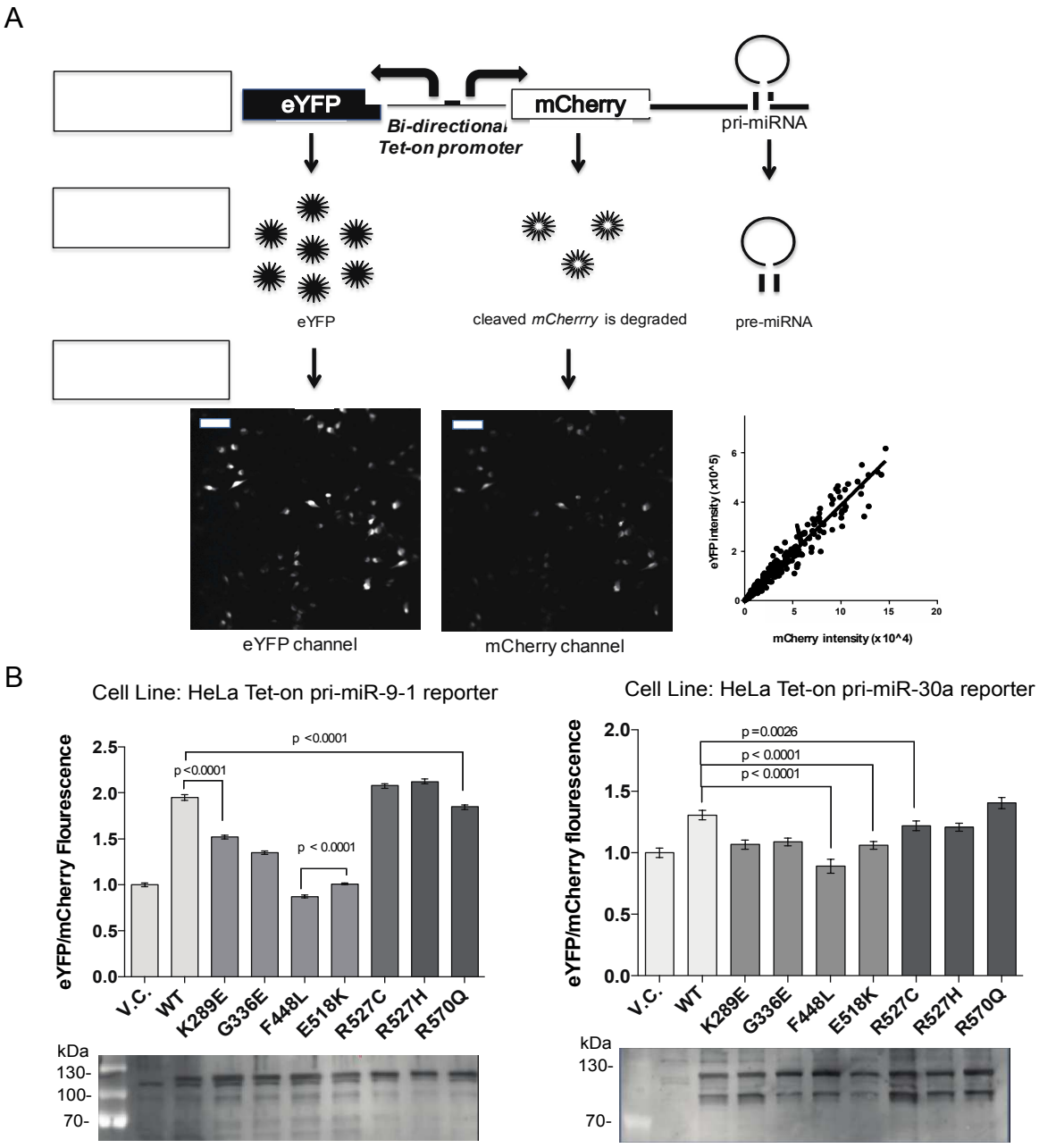


Figure 2. Processing activity of tumor-derived mutants in human cells. (A) Live-cell processing assay bidirectional reporter design for measurement of microRNA processing in cells. Scale bar on cell

images is 50 nm. (B) Processing activity of each DGCR8 mutant in both HeLa and U2-OS cells with immunoblots indicating expression of each mutant is near or above DGCR8 overexpression.

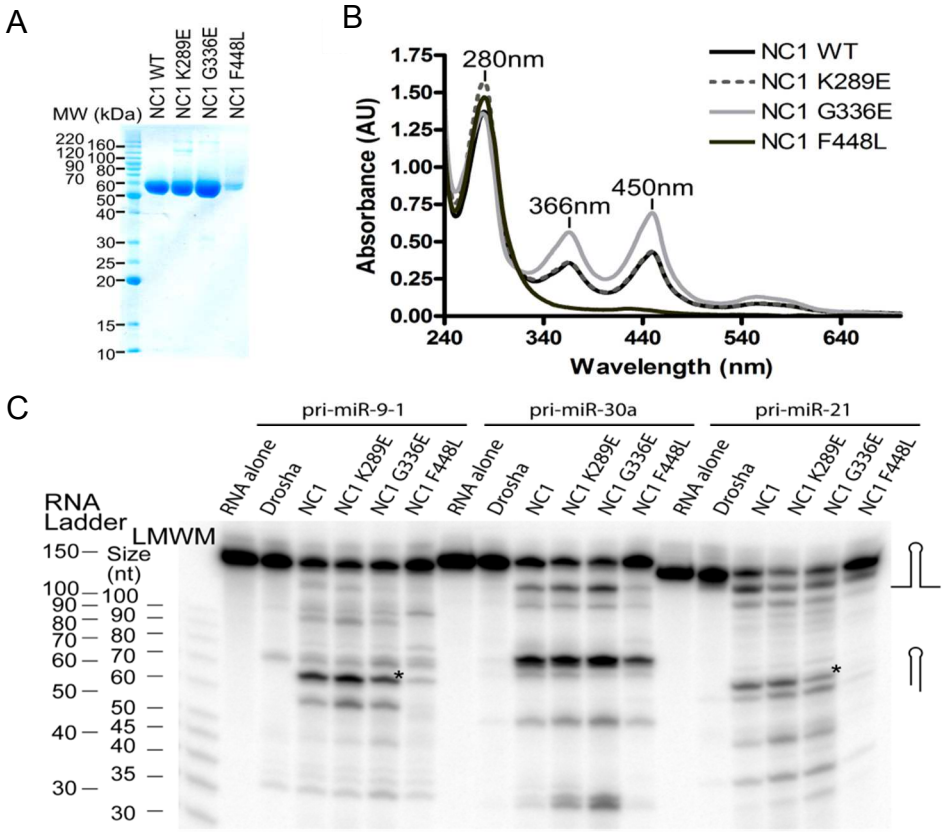


Figure 3. **Biochemical characterization of DGCR8 Rhed mutants.** (A) Coomassie -stained SDS gel of purified NC1 proteins. (B) Electronic absorption spectra of NC1 WT and mutant proteins. (C) In vitro Microprocessor assay comparing the three Rhed mutants across three pri-miRNA substrates. We previously determined the relative electrophoretic migration of the Low Molecular Weight Marker (LMWM, Affymetrix) and the RNA Decade ladder (ThermoFisher). Pri-miRNA substrates are 150 nt and pre-miRNA products are between 50-60 nt.

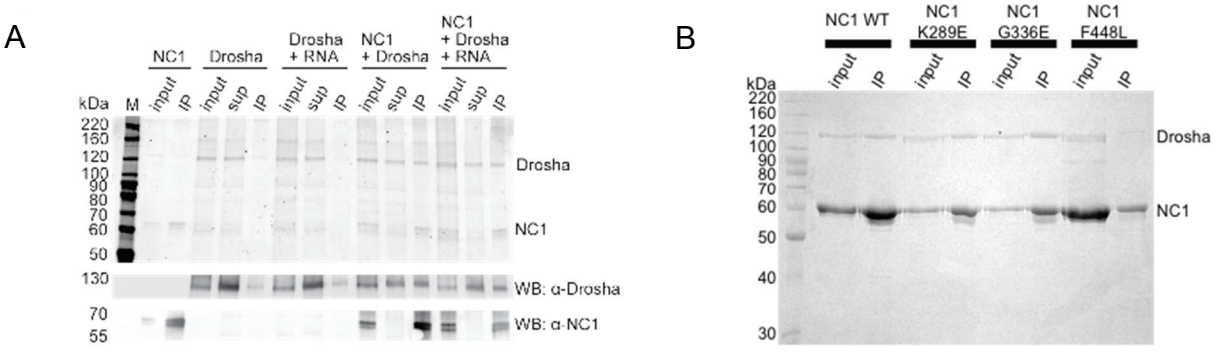


Figure 4. Drosha pull-down assay with N-FLAG-NC1. (A) Validation of the pull down experiment showing that Drosha selectively precipitates in the presence of N-FLAG-NC1. Top panel shows a SyproRed-stained SDS gel and lower panels are immunoblots for Drosha and NC1. (B) Coomassie-stain SDS gel of the pull-down assay for mutations in the Rhed.

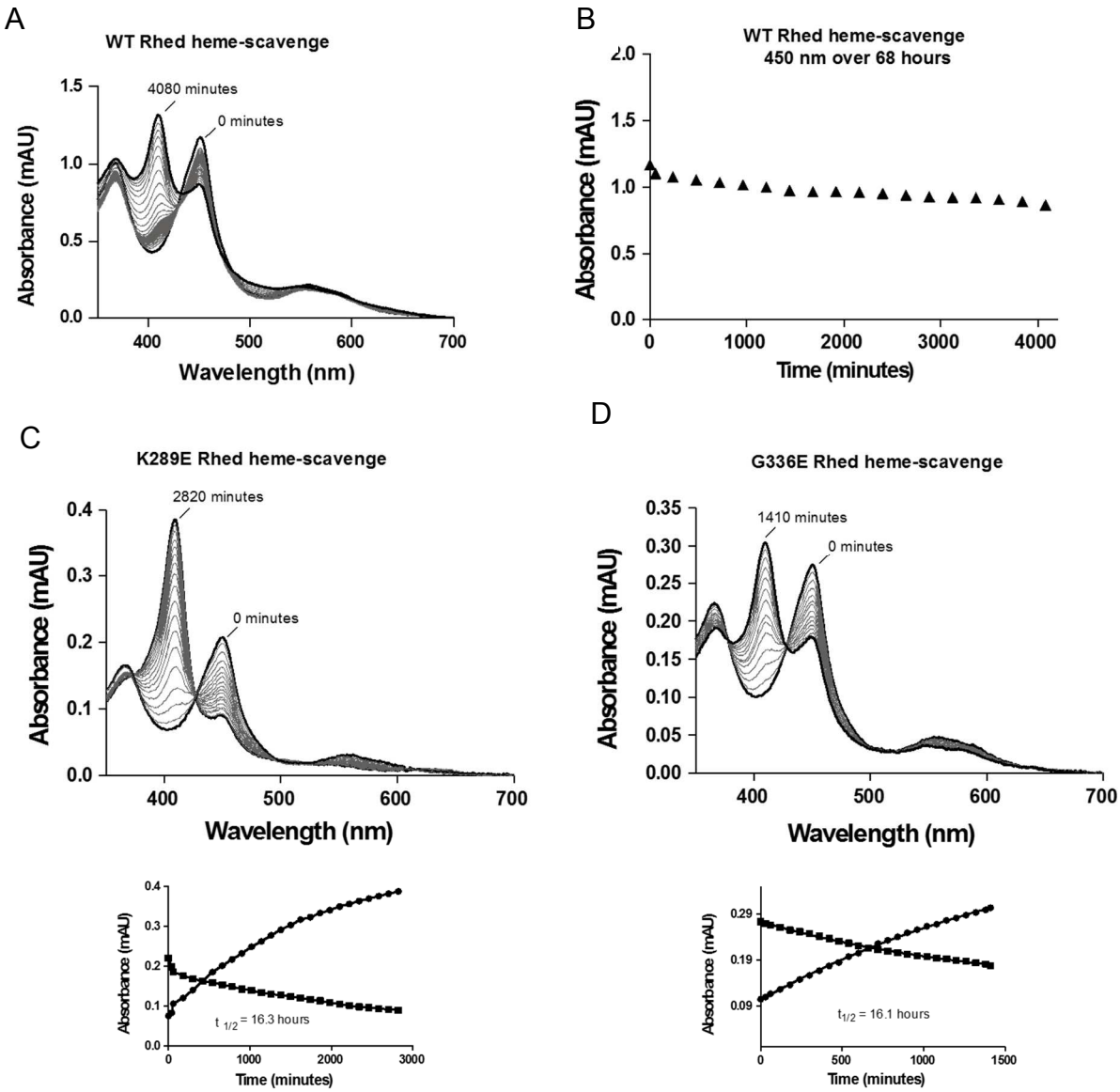


Figure 5. Tumor-derived point-mutations in the Rhed weaken DGCR8s binding affinity for heme. (A-B) Heme-scavenging by apomyoglobin is monitored by UV-VIS. (C-D) Plots of 409 nm absorbance from ferric myoglobin and concomitant decrease in 450 nm absorbance from DGCR8 söret peak Data is fit with one-phase decay to calculated the half-life of heme association with each mutant.

A

Protein	Kd reported in nM		
	pri-miR-9-1	pri-miR-21	pri-miR-30a
WT NC1	10 ± 1	12±4	16±1
K289E (NC1)	36 ±6	11±5	11±1
G336E (NC1)	52±14	49±14	11±2
K289E (Rhed)	no saturaton	0-2uM	
G336E (Rhed)	no saturation	0-2uM	

B

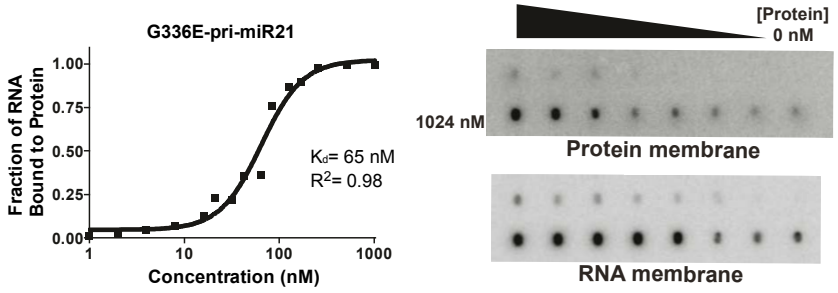


Figure 6. K289E and G336E have altered pri-miRNA binding properties. (A) Dissociation constants for K289E and G336E in DGCR8 with three different miRNAs. Previously published WT NCI K_d values are highlighted in grey. (B) Example curve and raw dotblot data. Data is fit with a cooperative dimer model.

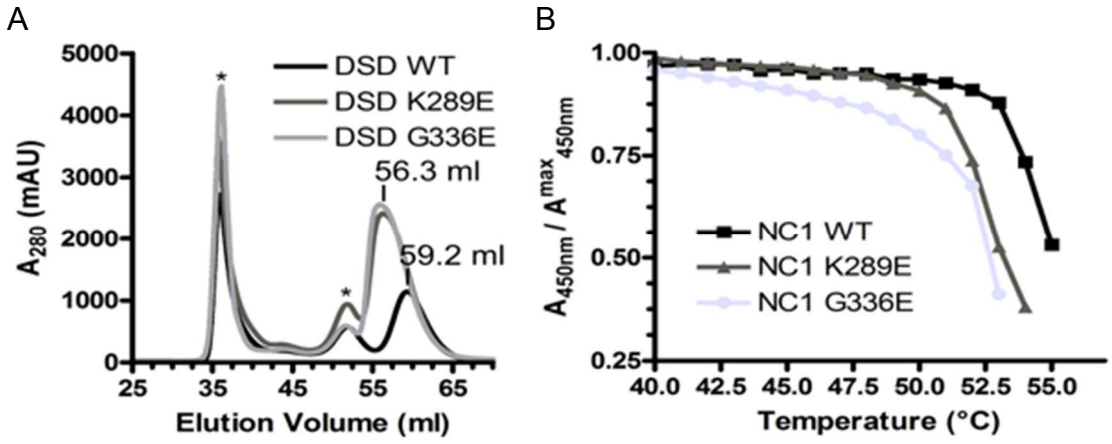


Figure 7. K289E and G336E have decreased stability compared to WT DGCR8. (A) Size-exclusion chromatography analysis of DSD mutant proteins. Aggregated and RNA-bound protein peaks are marked

with an asterisk. (B) Melting curves for NC1 mutant proteins analyzed at the 450 nm absorption peak of the heme-protein complex.

Materials and Methods

Live-cell pri-miRNA processing reporter assay

Tet-on Hela and U2-OS cells were grown at 37°C with 5% CO_2 in DMEM medium containing 4.5 g/L glucose, L-glutamine and 110 mg/L sodium pyruvate supplemented with 10% Tet-System-Approved FBS (Clonetech Laboratories). For imaging, cells were first passaged into 3.5-cm dishes with a glass coverslip bottom. Cells were transfected at ~80% confluency with Effectene (Qiagen), and the expression of reporters was immediately induced by addition of 1 µg/mL doxycycline. About 20 hr posttransfection, cells were washed and exchanged into DMEM without phenol red. Fluorescence imaging was performed using a Nikon Eclispse Ti microscope with an EMCCD camera (Andor iXon). We used a 10x objection lens and an exposure time of 0.3 ms with gain at 0. Fluorescence was detected with filter cubes for eYFP (excitation filter at 510 ± 10 nm band pass and emission filter at 535 ± 15 nm band pass) and mCherry (excitation filter at 535 ± 25 nm band pass and emission filter at 610 ± 10 nm band pass). Images were exported as 16-bit tiff files.

Images were analyzed mostly as described (4), except that a new Matlab (The MathWorks) script was written for integrating fluorescence intensities of individual cells in a streamlined fashion. Using the image analysis toolbox in Matlab, our program first performs segmentation to determine the locations of individual nuclei from pairs of eYFP and mCherry images acquired from the same field of view. A mask is created such that pixels with intensity above a user-defined threshold are considered to be in nuclei (both eYFP and mCherry expressed contain a nuclear localization sequence) and are assigned a value of 1. The values of all other pixels are set to 0. We choose threshold values that maximize the number of cells detected. The perimeter of a nucleus is determined for each image of a pair separately, and then a common mask is chosen for the two images based on the following criteria. If the same

cell is found in both images, the mask with smaller perimeter is chosen to avoid including cytosols, which sometimes have low fluorescent signals. If a cell is detected in only one image, the perimeter is kept, allowing for cells that are dim in the other channel to be included. The center of the final perimeter for each cell is determined. A background fluorescence value is calculated for each image by averaging the intensities of the pixels outside the identified nuclei. The background is then subtracted from the intensities of all pixels. The total intensity of an individual cell is calculated as the sum of the pixel intensities within the perimeter. Dead cells often have very strong fluorescent signals, therefore cells with top 10% fluorescent intensities are excluded in the analyses. Overlapping cells that are detected as one object and can be removed manually. A list of total eYFP and mCherry intensities and their center coordinates for individual cells is exported to Microsoft Excel, scatter plots are generated and linear fit of eYFP versus mCherry signals, forced to go through the origin, gives the fluorescence slopes.

DGCR8 bacterial expression and purification

NC1, Rhed, and DSD, including WT and tumor-derived mutants, were expressed and purified as previously reported (5). Briefly, DGCR8 proteins were expressed in *E. coli* strain BL21(DE3). Cells were grown to late-log phase. Expression was induced with 1 mM IPTG for 4 hr at 37°C. For NC1 and Rhed, δ-aminolevulinic acid was added to a final concentration of 1 mM to facilitate heme loading to the proteins. Cells were harvested by centrifugation at 5000xg, for 5 min at 4°C. Protein was initially purified using a SP cation-exchange column (GE Healthcare). Column equilibration and loading were performed using Buffer *A*, which contained 100 mM NaCl and 20 mM Tris pH 8.0. Linear gradient elution was performed with Buffer B containing 2 M NaCl and 20 mM Tris pH 8.0. Fresh dithiothreitol (DTT) was added to buffers at a final concentration of 1 mM. Fractions containing the DGCR8 protein were concentrated in a centrifugal concentrator with appropriate MWCO (Millipore), and loaded onto a Superdex-200 size exclusion chromatography column (GE Healthcare) equilibrated with a SEC buffer

containing 400 mM NaCl and 20 mM Tris pH 8.0. Peak fractions were again concentrated and analyzed using a SDS 12%-polyacrylamide gel to determine purity.

Pri-miRNA binding and processing assays

The pri-miRNA binding assays were performed as described previously (6). Briefly, serial dilutions of purified DGCR8 protein were incubated with a trace amount of pri-miRNA uniformly labeled with ^{32}P in 100 mM NaCl and 20 mM Tris pH 8.0 at room temperature for 30 min. The mixtures were filtered through stacked nitrocellulose (EMD Millipore) and positively changed nylon (GE Healthcare) membranes. The autoradiography images of the membranes were analyzed using Quantity One (Bio-Rad, version 4.4.1). The data were fit using PRISM (GraphPad, version 6).

Pri-miRNA processing assays were performed as described (5).

Heme dissociation assay

Purified DGCR8 protein at 3-20 μM concentration was incubated in SEC buffer at room temperature with 4x or 5x molar excess of horse skeletal apomyoglobin (SIGMA). Electronic absorption spectra were monitored over time using a double-beam Varian CARY 300 Bio spectrophotometer with spectral bandwidth set to 1.0 nm, and averaging time at 0.1 sec. The scanning kinetics program was used to automatically collect scans at set intervals. The data was fit with a one-phase exponential decay equation using PRISM.

References

1. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Europe PMC Funders Group Patterns of somatic mutation in human cancer genomes. 2009;446(7132):153–8.

2. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. 2000;21(3):379–85.
3. Stoler DL, Chen N, Basik M, Kahlenberg MS, Rodriguez-bigas MA, Petrelli NJ, et al. The onset and extent of genomic instability in sporadic colorectal tumor progression. 1999;

4. Endesfelder D, Math D, Gronroos E, Ph D, Martinez P, Ph D, et al. new england journal. 2012;

5. Stratton MR. Exploring the Genomes of Cancer Cells : Progress and Promise. 2011;(March):1553–9.

6. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. 2015;112(1).

7. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. 2007;39(5):673–7.

8. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435(7043):834–8.

9. Almeida MI, Reis RM, Calin G a. MicroRNA history: discovery, recent applications, and next frontiers. Mutat Res. Elsevier B.V.; 2011; 717(1–2):1–8.

10. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, et al. microRNAs exhibit high frequency genomic alterations in human cancer.

11. Zhang B, Pan X, Cobb GP, Anderson T a. microRNAs as oncogenes and tumor suppressors. Dev Biol. 2007;302(1):1–12.

12. Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, et al. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. Cell [Internet]. Elsevier Inc.; 2013 Jan 17 [cited 2014 Jul 15];152(1–2):276–89.

13. Chivukula R, Shi G, Acharya A, Mills E, Zeitels LR, Anandam JL, et al. An essential mesenchymal function for miR-143/145 in intestinal epithelial regeneration. Cell. 2014;157(5):1104–16.

14. Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Heerema N, et al. Pre-B cell proliferation and lymphoblastic leukemia ¯high-grade lymphoma in E □
-miR155 transgenic mice. 2006;103(18):2–7.

15. Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J, et al. Dicer1 functions as a haploinsufficient tumor suppressor. Genes Dev. 2009;23(23):2700–4.

16. Lee Y, Jeon K, Lee J, Kim S, Kim VN. MicroRNA maturation : stepwise processing and subcellular localization. 2002;21(17):4663–70.

17. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. 2009;10.

18. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004;116(2):281–97.

19. Merritt WM, Lin YG, Han LY, Kamat AA, Spannuth WA, Schmandt R, et al. Dicer, Drosha, and Outcomes in Patients with Ovarian Cancer.

20. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nat Genet. 2007;39(5):673–7.

21. Bandara K V, Michael MZ, Gleadle JM. Hypoxia represses microRNA biogenesis proteins in breast cancer cells. BMC Cancer. 2014;14(1):533.

22. Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, et al. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell . NIH Public Access; 2015 Feb 9;27(2):286–97.

23. Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M, et al. Mutations in the SIX1/2 Pathway and the DROSHA/DGCR8 miRNA Microprocessor Complex Underlie High-Risk Blastemal Type Wilms Tumors. Cancer Cell [Internet]. 2015 Feb 9;27(2):298–311.

24. Belair CD, Paikari A, Moltzahn F, Shenoy A, Yau C,Era M, et al. DGCR8 is essential for tumor progression following PTEN loss in the prostate. EMBO Rep. 2015;16(9):1219–32.

25. Quick-Cleveland J, Jacob JP, Weitz SH, Shoffner G, Senturia R, Guo F. The DGCR8 RNA-Binding Heme Domain Recognizes Primary MicroRNAs by Clamping the Hairpin. Cell Rep. 2014 Jun 26 ;7(6):1994–2005.

26. Nguyen TA, Jo MH, Choi Y-G, Park J, Kwon SC, Hohng S, et al. Functional Anatomy of the Human Microprocessor. Cell [Internet]. Elsevier Inc.; 2015;161(6):1374–87.

27. Faller M, Matsunaga M, Yin S, Loo J a, Guo F. Heme is involved in microRNA processing. Nat Struct Mol Biol [Internet]. 2007;14(1):23–9.

28. Barr I, Smith AT, Chen Y, Senturia R, Burstyn JN, Guo F. Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. Proc Natl Acad Sci U S A. 2012 Feb 7;109(6):1919–24.

29. Herbert KM, Pimienta G, DeGregorio SJ, Alexandrov A, Steitz JA. Phosphorylation of DGCR8 Increases Its Intracellular Stability and Induces a Progrowth miRNA Profile. Cell Rep; 2013;5(4):1070–81.

30. Zhu C, Chen C, Huang J, Zhang H, Zhao X, Deng R, et al. SUMOylation at K^{707} of DGCR8 controls direct function of primary microRNA. Nucleic Acids Res. 2015;43(16):7945–60.

31. Wada T, Kikuchi J, Furukawa Y. Histone deacetylase 1 enhances microRNA processing via deacetylation of DGCR8. Nat Publ Gr. Nature Publishing Group; 2012;13(2):142–9.

32. Gong M, Chen Y, Senturia R, Ulgherait M, Faller M, Guo F. Caspases cleave and inhibit the microRNA processing protein DiGeorge Critical Region 8. Protein Sci.

2012;21(6):797–808.

33. Han J, Pedersen JS, Kwon SC, Belair CD, Kim Y, Yeom K, et al. Posttranscriptional Crossregulation between Drosha and DGCR8. Cell [Internet]. Elsevier Inc.; 2009;136(1):75–84.

34. Triboulet R, Chang H-M, Lapierre RJ, Gregory RI. Post-transcriptional control of DGCR8 expression by the Microprocessor. RNA [Internet]. 2009;15(6):1005–11.

35. Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, et al. Sequestration of DROSHA and DGCR8 by expanded CGG RNA Repeats Alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep. The Authors; 2013;3(3):869–80.

36. Senturia R, Faller M, Yin S, Loo J a., Cascio D, Sawaya MR, et al. Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. 2010;19(7):1354–65.

37. Senturia R, Laganowsky A, Barr I, Scheidemantle BD, Guo F. Dimerization and heme binding are conserved in amphibian and starfish homologues of the microrna processing protein DGCR8. PLoS One. 2012;7(7).

38. Sun Young Sohn, Won Jin Bae, Jeong Joo Kim, Kyu-Hyeon Yeom, V. Narry Kim and Yunje Cho. Crystal structure of human DGCR8 core. Nature Structural & Molecular Biology.2007;14(9):847-853.

39. Weitz SH, Gong M, Barr I, Weiss S, Guo F. Processing of microRNA primary transcripts requires heme in mammalian cells. Proc Natl Acad Sci U S A. 2014;111(5):1861–6.

40. Kwon SC, Nguyen TA, Choi YG, Jo MH, Hohng S, Kim VN, et al. Structure of Human DROSHA. Cell; 2016;164(1–2):81–90.

41. Herbert KM, Sarkar SK, Mills M, Delgado De la Herran HC, Neuman KC, Steitz JA. A heterotrimer model of the complete Microprocessor complex revealed by single-molecule subunit counting. RNA. 2016;22(2):175–83.

42. Han J, Lee Y, Yeom K, Kim Y, Jin H, Kim VN. The Drosha – DGCR8 complex in primary microRNA processing. Gene & Development.2004;1–12.

43. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the BRAF gene in human cancer. 2002;949–54.

44. Kebebew E, Bauer J, Clark OH, Shibru D, Bastian B, Griffin A. The Prevalence and Prognostic Value of BRAF Mutation in. 2007;246(3):466–71.

45. Ascierto PA, Kirkwood JM, Grob J, Simeone E, Grimaldi AM, Maio M, et al. The role of BRAF V600 mutation in melanoma. 2012;1–9.

46. Holderfeild M, Deuker M, McCormick F, McMahon M. Targeting RAF kinases for cancer therapy: BRAF mutated melanoma and beyond. Nat Rev Cancer. 2014;14(7):455–67.

47. Teresa de la Morena. M, Eitson, Jennifer L., Dozmorov, Igor M., Belkaya, Serkan, Hoover, Ashley R., Anguiano, Esperanza, Pascual, M. Virginia, van Oers NSC. Signature MicroRNA expression patters identified in humans with 22q11.2 deletion/DiGeorge

syndrome. Clin Immunol. 2013;147(1):11–22.

48. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2002;99(24):15524–9.

49. Mori M, Triboulet R, Mohseni M, Schlegelmilch K, Shrestha K, Camargo FD, et al. NIH Public Access. 2015;156(5):893–906.

50. Barr I, Weitz SH, Atkin T, Hsu P, Karayiorgou M, Joseph A. Cobalt (III) protoporphyrin activates the DGCR8 protein and can compensate microRNA processing deficiency. Chemistry & Biology. 2015; 22 (6):793-802.

51. Chang TC, Pertea M, Lee S, Salzberg SL, Mendell JT. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. Genome Res. 2015;25(9):1401–9.

Appendix A: Jose2dCellCounter and Live-Cell Assay

About the program

This program was written by Dr. Jose Rodriguez based on original code by Dr. Sara Weitz. I beta tested the early versions of the program and helped develop it into its current form.

Jose2dCellCounter works by integrating fluorescence intensities of individual cells. Using the image analysis toolbox in Matlab, the program first performs segmentation to determine the locations of individual nuclei from pairs of eYFP and mCherry images acquired from the same field of view. A mask is created such that pixels with intensity above a user-defined threshold are considered to be in nuclei (both eYFP and mCherry expressed contain a nuclear localization sequence) and are assigned a value of 1. The values of all other pixels are set to 0. We choose threshold values that maximize the number of cells detected. The perimeter of a nucleus is determined for each image of a pair separately, and then a common mask is chosen for the two images based on the following criteria. If the same cell is found in both images, the mask with smaller perimeter is chosen to avoid including cytosols, which sometimes have low fluorescent signals. If a cell is detected in only one image, the perimeter is kept, allowing for cells that are dim in the other channel to be included. The center of the final perimeter for each cell is determined. A background fluorescence value is calculated for each image by averaging the intensities of the pixels outside the identified nuclei. The background is then subtracted from the intensities of all pixels. The total intensity of an individual cell is calculated as the sum of the pixel intensities within the perimeter. Both dead and overlapping cells that are detected as one object and can be removed manually. A list of total eYFP and mCherry intensities and their center coordinates for individual cells is exported to as csv from which scatter plots are generated. Linear fitting of eYFP versus mCherry signals, forced to go through the origin, gives the fluorescence slopes.

Jose2dCellCounter code

```
function [count, intensities, totalcells]=JOSE_2dcellcounter(nimages,guiswitch)

if exist('guiswitch','var')==0
    guiswitch='off';
end
folder_name = uigetdir();
old_wd=cd(folder_name);
osy=100; osx=100; sy=1600; sx=800;

warning('off','all');
count=zeros(1,nimages,'single');
intensities=cell(nimages,1); %count;
coords=intensities;
half_small_box=6;
half_big_box=3*half_small_box+1;

cin=single(imread(strcat('yfp_1.tif')));
if size(cin,3)>1
    cin=cin(:,:,1);
end
[Y X]=size(cin);
halfx=floor(X/2)+1;
halfy=floor(Y/2)+1;

fwhmy=Y/4;
fwhmx=X/4;
a= 1:1:X;
b= 1:1:Y;
[aa,bb]=meshgrid(a,b);
kfilter=exp( - ( ((aa-halfx).^2)./(2*fwhmx^2) + ((bb-halfy).^2)./(2*fwhmy^2) ) );
kfilter=kfilter/max(max(kfilter));
clear a b aa bb
kk=1; backone=0; sval=1.5;

while kk<=nimages

    c1o=single(imread(strcat('mch_',num2str(kk),'.tif')));

    if size(c1o,3)>1
        c1o=squeeze(c1o(:,:,1));
    end
    c1=real(ifftn(ifftshift(fftshift(fftn(c1o)).*kfilter)));

    c2o=single(imread(strcat('yfp_',num2str(kk),'.tif')));
    %whos;
    if size(c2o,3)>1
        c2o=squeeze(c2o(:,:,1));
    end
    whos;
    %figure, subplot(1,2,1), imagesc(c2o), colormap gray;
    c2=real(ifftn(ifftshift(fftshift(fftn(c2o)).*kfilter)));
    %subplot(1,2,2), imagesc(c2), colormap gray;
```

```
    c1o_lin=c1o(c2o>0); size(c1o_lin),
    c2o_lin=c2o(c2o>0); size(c2o_lin),
    %c1o_mn=mean(c1o_lin);
    c2o_mn=mean(c2o_lin);
    %c1o_sd=std(c1o_lin);
    c2o_sd=std(c2o_lin);

    c1o_bg=c1o_lin(c2o_lin<(c2o_mn+sval*c2o_sd));

   c1o_cs=c1o_lin(c2o_lin>=(c2o_mn+sval*c2o_sd));
    c2o_bg=c2o_lin(c2o_lin<(c2o_mn+sval*c2o_sd));
    c2o_cs=c2o_lin(c2o_lin>=(c2o_mn+sval*c2o_sd));

    hi=figure(9);
    subplot(1,3,1), plot(c1o_lin,c2o_lin,'.k'), axis([min(c1o_lin) max(c1o_lin) min(c2o_lin) max(c2o_lin)]);
    xlabel(['MCH ',int2str(kk),' all points']), ylabel(['YFP ',int2str(kk),' all points']), title(strcat(['Image #',int2str(kk),'all']));
    subplot(1,3,2), plot(c1o_bg,c2o_bg,'.k'), axis([min(c1o_lin) max(c1o_lin) min(c2o_lin) max(c2o_lin)]);
    xlabel(['MCH ',int2str(kk),' all points']), ylabel(['YFP ',int2str(kk),' all points']), title(strcat(['Image #',int2str(kk),'bg']));
    subplot(1,3,3), plot(c1o_cs,c2o_cs,'.k'), axis([min(c1o_lin) max(c1o_lin) min(c2o_lin) max(c2o_lin)]);
    xlabel(['MCH ',int2str(kk),' all points']), ylabel(['YFP ',int2str(kk),' all points']), title(strcat(['Image
#',int2str(kk),'cells']));

    if strcmp(guiswitch,'gui')==0
        cin=real(ifftn(ifftshift(fftshift(fftn(c2o)).*kfilter)));
        cmean=mean(cin(cin>0));
        cstd=std(cin(cin>0));
        snr=floor(((cmean+sval*cstd)/cmean)*100)/100; %snr=1.1;
    else
        hf=figure(11);
        set(hf, 'Position', [osy, osx, sy, sx]);
        subplot(1,2,1:2), imagesc(c2o), axis image, title('Click on a representative cell','FontSize',14), colormap gray;
        [xin,yin]=ginput(1);
        cell_int=sum(sum(c2o(yin-half_small_box:yin+half_small_box,xin-half_small_box:xin+half_small_box)));

        hf=figure(11);
        set(hf, 'Position', [osy, osx, sy, sx]);
        subplot(1,2,1:2), imagesc(c2o), axis image, title('Click on a clear background area','FontSize',14), colormap
gray;
        [xin,yin]=ginput(1);
        delete(hf);
        bg_int=sum(sum(c2o(yin-half_small_box:yin+half_small_box,xin-half_small_box:xin+half_small_box)));
        snr=floor((cell_int/bg_int)*100)/100;
    end

    mean1=mean(mean(c1));
    mean2=mean(mean(c2));

    lp=1;
    while lp==1

    c1_snr=c1/mean1; c1_bin=c1_snr;
    c2_snr=c2/mean2; c2_bin=c2_snr;

    c_snr=max(c1_snr,c2_snr);

    c1_bin(c1_bin<snr)=0; c1_bin(c1_bin>snr)=1;
    c2_bin(c2_bin<snr)=0; c2_bin(c2_bin>snr)=1;

    se=strel('square',4);
    c_bin=max(c1_bin,c2_bin);
    [y1 x1]=find(c_bin==1);

    numpix=max(size(y1));
```

```
    sbox_mask=zeros(Y,X,'single');
    bbox_mask=sbox_mask;

    for jj=1:numpix
        if y1(jj)-half_big_box>0 && x1(jj)-half_big_box>0 && y1(jj)+half_big_box<Y && x1(jj)+half_big_box<X
        sbox_mask(y1(jj),x1(jj))=sum(sum(c_snr(y1(jj)-half_small_box:y1(jj)+half_small_box,...
            x1(jj)-half_small_box:x1(jj)+half_small_box)));
        bbox_mask(y1(jj),x1(jj))=sum(sum(c_snr(y1(jj)-half_big_box:y1(jj)+half_big_box,...
            x1(jj)-half_big_box:x1(jj)+half_big_box)));
        end
    end
    bbox_mean=(bbox_mask-sbox_mask)./(half_big_box+1)^2;
    sbox_mean=sbox_mask./(half_small_box+1)^2;
    rat_mask=max(1,sbox_mean./bbox_mean);

    cutoff=snr;
    cmask=rat_mask; cmask(cmask<cutoff)=0; cmask(cmask>=cutoff)=1;
    cmask=imerode(cmask,se);
    cmask=imdilate(cmask,se);
    cmask=imfill(cmask,'holes');


    cells_count= bwlabel(cmask);
    cells_outline = bwperim(cmask);
    c1f=c1o; c1f(cells_outline>0)=max(max(c1f));
    c2f=c2o; c2f(cells_outline>0)=max(max(c2f));

    hf=figure(11);
    set(hf, 'Position', [osy, osx, sy, sx]);
    subplot(1,2,1), imagesc(c1f), axis image, title(['Youre on image',num2str(kk),'with SNR=
',num2str(snr)],'FontSize',14);
    subplot(1,2,2), imagesc(c2f), axis image, title('Press space to continue, up or down arrows to change SNR
level','FontSize',14);
    drawnow(); colormap gray;
    [~,~,button]=ginput(1);

    if button==30 % 30= up arrow, larger SNR
        snr=snr+0.01;
    elseif button==29 % 29= right arrow, much higher SNR
        snr=snr+0.1;
    elseif button==31 % 31= down arrow, lower SNR
        snr=snr-0.01;
    elseif button==28 % 28= left arrow, much lower SNR
        snr=snr-0.1;
    else
        lp=0;
    end

    end

    hf=figure(11);
    set(hf, 'Position', [osy, osx, sy, sx]);
    subplot(1,2,1), imagesc(c2o), axis image, title(['Original Image',num2str(kk),'with SNR=
',num2str(snr)],'FontSize',14);
    subplot(1,2,2), imagesc(c2f), axis image, title('Click on the dead / double cells HERE, then press
ENTER','FontSize',14);
    drawnow(); colormap gray;
    [xdead,ydead]=ginput();
    ndead=max(size(ydead));

    cells_count2=cells_count;
    if ndead>=1
        for mm=1:ndead
```

```
            yd=round(ydead(mm));
            xd=round(xdead(mm));
            cells_count2(cells_count==cells_count(yd,xd))=0;
        end
    end

    cmask=cells_count2; cmask(cmask>0)=1;
    cells_outline = bwperim(cmask);
    c2f=c2o; c2f(cells_outline>0)=max(max(c2f));

    hf=figure(11);
    set(hf, 'Position', [osy, osx, sy, sx]);
    subplot(1,2,1), imagesc(c2o), axis image, title(['Original Image:',num2str(kk),'with SNR=
',num2str(snr)],'FontSize',14);
    subplot(1,2,2), imagesc(c2f), axis image, title('Click to revive cells, then press ENTER','FontSize',14);
    drawnow(); colormap gray;
    [xrelive,yrelive]=ginput();
    nrelive=max(size(yrelive));

    if nrelive>=1
        for mm=1:nrelive
            yd=round(yrelive(mm));
            xd=round(xrelive(mm));
            cells_count2(cells_count==cells_count(yd,xd))=...
                cells_count(cells_count==cells_count(yd,xd));
        end

    cmask=cells_count2; cmask(cmask>0)=1;
    cells_outline = bwperim(cmask);
    c1f=c1o; c1f(cells_outline>0)=max(max(c1f));
    c2f=c2o; c2f(cells_outline>0)=max(max(c2f));
    end

    hf=figure(11);
    set(hf, 'Position', [osy, osx, sy, sx]);
    subplot(1,2,1), imagesc(c2o), axis image, title(['Original Image:',num2str(kk),' Gui status:',guiswitch],'FontSize',14);
    subplot(1,2,2), imagesc(c2f), axis image, title('These are your final cells! type g for gui, b for back, then press
ENTER','FontSize',14);
    drawnow(); colormap gray;
    [~,~,button]=ginput();
    delete(hf);
    nfinal=max(size(button));

    if nfinal>=1
        for mm=1:nfinal
            if button(mm)==98 % 98=b, back
                backone=1;
            elseif button(mm)==103 && strcmp(guiswitch,'gui')==0 %103=g, gui
                guiswitch='gui';
            elseif button(mm)==103
                guiswitch='off';
            end
        end
    end

    count(kk)=max(max(cells_count));
    ints1=[]; ints2=[]; ycoords=[]; xcoords=[];

    c1bg=min(c1o(cmask>0));
    c2bg=min(c1o(cmask>0));
    c1c=max(0,c1o-c1bg);
    c2c=max(0,c2o-c2bg);
    desc=regionprops(cells_count2,'Centroid');
```

```
    ctt=0;
    for mm=1:count(kk)
        if sum(sum(cells_count2(cells_count2==mm)))>1
            ctt=ctt+1;
            lcoords=desc(mm).Centroid;
            ints1(ctt)=sum(sum(c1c(cells_count2==mm)));
            ints2(ctt)=sum(sum(c2c(cells_count2==mm)));
            ycoords(ctt)=round(lcoords(1));
            xcoords(ctt)=round(lcoords(2));
        end
    end
    intensities{kk,1,:}=ints1;
    intensities{kk,2,:}=ints2;
    coords{kk,1,:}=ycoords;
    coords{kk,2,:}=xcoords;
    imwrite(uint16(c1f),strcat('mch_',num2str(kk),'o.tif'),'tif');
    imwrite(uint16(c2f),strcat('yfp_',num2str(kk),'o.tif'),'tif');

    kk=kk+1-backone;
    backone=0;
end

images_wd=cd(old_wd);
pathname=textscan(images_wd,'%s','delimiter','/');
lastname=max(size(pathname{1}));
namecore=strcat(pathname{1}{lastname-1},'_',pathname{1}{lastname});
old_wd=cd(images_wd);

%a1=floor(sqrt(nimages)); a2=ceil(nimages/a1);
ha=figure(12);
hold off;
set(ha, 'Position', [25, 50, 500, 400]);

fid=fopen(strcat(namecore,'integ.csv'),'w');
for kk=1:nimages

    s1=max(size(intensities{kk,1,:}));
    if kk==1
        cind1=1;
    else
        cind1=max(size(allints1));
    end
    allints1(cind1+1:cind1+s1)=intensities{kk,1,:};
    allints2(cind1+1:cind1+s1)=intensities{kk,2,:};

    numColumns=max(size(intensities{kk,1,:}));

    for jj=1:numColumns
        if jj==1
            fwrite(fid,strcat('Image_',num2str(kk),'_mch',','),'uchar');
            fwrite(fid,strcat('Image_',num2str(kk),'_yfp',','),'uchar');
            fwrite(fid,strcat('Image_',num2str(kk),'_yc',','),'uchar');
            fwrite(fid,strcat('Image_',num2str(kk),'_xc',','),'uchar');
            fprintf(fid,'\n');
        end
        fprintf(fid,'%f,',intensities{kk,1,:}(jj));
        fprintf(fid,'%f,',intensities{kk,2,:}(jj));
        fprintf(fid,'%f,',coords{kk,1,:}(jj));
        fprintf(fid,'%f,',coords{kk,2,:}(jj)); fprintf(fid,'\n');
    end

    plot(intensities{kk,1,:},intensities{kk,2,:},' o','MarkerEdgeColor',[0 (kk/nimages) 1-(kk/nimages)],...
        'MarkerFaceColor',[0 (kk/nimages) 1-(kk/nimages)],'MarkerSize',4); %subplot(a1,a2,kk),
```

```
    xlabel(['MCH ',int2str(kk)]), ylabel(['YFP ',int2str(kk)]), title(['Ratios for all ',int2str(kk),' images']);
    axis([0 max(allints2) 0 max(allints2)]);
    hold on;
end

fclose(fid);
saveas(ha,strcat(namecore,'_plots.png'),'png');
saveas(hi,strcat(namecore,'_scatter_all.png'),'png');

totalcells=max(size(allints1));
p = polyfit(allints1,allints2,1);
yfit = polyval(p,allints1);
yresid = allints2 - yfit;
SSresid = sum(yresid.^2);
SStotal = (length(allints2)-1) * var(allints2);
rsq = 1 - SSresid/SStotal;
fprintf(strcat('The equation: YFP = ',num2str(p(1)),'*MCH + ',num2str(p(2)),'has an Rsquared value of ',num2str(rsq),
'\n'));

figure (15), plot(allints1,allints2,' o','MarkerEdgeColor','k',...
        'MarkerFaceColor','k','MarkerSize',4), axis([0 max(allints2) 0 max(allints2)]);
cd(old_wd);
```

Modified Procedure: Live-Cell Assay

Live-Cell Assay ~ original protocol by Sara Weitz, modified by Jen Quick-Cleveland

Day 1: Prepare cells for transfection. Cells should be split into glass bottom 2cm dishes for imaging. Cells should be ~80% confluent on the day of transfection.

Day 2: Transfection: We prefer effectene. Below is a transfection set –up table. Any transfection reagent that works is fine. DNA concentration can be varied but the DNA: enhancer ratio needs to be 1:8. Induce reporter expression with 1mg/mL doxycycline and be prepared to image 18-24 hours afterwards. FPs will begin expressing in ~8

hours and get to be bright enough for good images around 16 hours.

Constructs	conc. (ng/ul)	Vol. to add for 1 ug total DNA		
ptreBi 9-1	214	4.7		
ptreBi 9-1 M1	219	4.6		
Transfection set-up (ul)	reporter 9-1 (4 plates)	reporter 9-1 M1 (4 plates)		
EC buffer	93.6	93.7		
reporter in ptreBi	4.7	4.6		
Enhancer	8 ul	8 ul		
Effectene - add to each aliquot	12 ul	12 ul		
Total Volume/transfection	**120**	**120**		
Induce the reporter with doxicyclin 1mg/ml stock	2.1 ul	2.1		
Wash plates well with 1x DPBS before transfection				
Transfection reaction preparation:				
Mix DNA+EC buffer				
Add the enhancer -- pipette up and down to mix and vortex tor 1 seconds -- 5 minute RT incubation				
Add the effectene-- pipette up and down to mix and gently vortex for 10 seconds -- 10 minute RT incubation				
add 600ul of warmed media to each transfection set-up				
pipette up and down to mix				
add transfection mixture DROPWISE to imaging plate of 3X washed cells				
swirl plates in a circular motion to mix				
Let cells incubate with tranfection complex for 10-15 minutes in the incubator before induction (if possible - not required)				

Day 3: Image transfected cells

Before imaging, wash the cells 3 times with 1x DPBS (-/-) to remove any dead cells

Add 2 mls of fresh 1x DPBS (room temperature)

Turn on the Nikon eclipse Ti lamp and camera (in room 2018-Young Hall)

On button is one the bottom right

Turn on joystick stage

Two boxes for lamp and LED on upper shelf – turn them on

Change the objective to 10x

Hit objective button on the bottom left to move the objective

The front knob changes the magnification -→ put this at 1x in a zoomed-out position

We use two filter cubes: light green for tx red 530/550 and cyan/turquoise for yfp 520/540

the filter they are in the door bottom left side of the microscope

On the table to your right there is a shutter box to control light intensity 1= highest intensity

Turn on camera and or ccd ---it needs to cool down to -80 degrees

Start the Andor Solis software on computer

To set the exposure time: hit remote control icon and set exposure to 0.3s for a 300ms

View cells with white light first then switch to lamp and filters do a course focus

Turn of white light and go to the YFP channel. Open shutter to check cell fluorescence

To switch the view to the computer and press video icon to activator camera

Definition of a “good” image: more than 10 cells, variety of fluorescent intensities, cells spread out if possible, cells bright enough that the background is dark. If you area doesn’t meet these criteria your should keep looking.

find cell group and pick one cells to get in focus until you can see the nucleus

IMAGE YFP first because it is brighter and easier to focus on

Take picture by clicking the camera icon and turn off light with shutter to protect FPs from bleaching

Save image pairs as yfp_# of image or mch_# of image

Change filter to tx RED, focus on the same cells as before and take same picture of the txREd channel. Focus doesn’t usually need to be changed and if it does it is a very small adjustment

Go back to video mode for the next frame

Always move the stage in the same direction (example: start imaging at the top right and move left and down) to avoid imaging the same cells

Image at least 200 cells.

After imaging go back to eye view (top button on front of microscope)

Cells to biohazard waste, and prep the next plate for imaging

Covert and export images:

File→ batch conversion → E → Sara→ HBD→ reporter only→ok

File =sif

Select all

Convert to tiff 16 bit 0-65535

Analyze Tiff images with Jose 2d Cell counter

Appendix B: Massively-Parallel Pri-miRNA Processing in Live Cells

One of the greatest challenges in the miRNA field is trying to understand the phenotypic consequences of individual miRNA in a global fashion. With few exceptions, many studies are based on either a small representative group of pri-miRNA or individual hairpins. Beyond this, the observable effect of changes in miRNA expression are notoriously subtle. Many of the broader studies on big groups of miRNA were done with miRNA arrays. With the emergence of NGS, looking at miRNA has become easier using methods like RNA-sequencing and Hi-throughput sequencing and Cross-linked Immunoprecipitation (HITS-CLIP).

Can we design a method to measure pri-miRNA processing in a global scale?

. Looking at miRNA as a regulatory *population* may be a more realistic approach to understanding their physiological effect. If techniques are developed to measure changes in miRNA production as a whole, I fully expect clear patterns to emerge. I speculate that these patterns will be governed by developmental time, external stimuli (infection, stress), and tissue type. Further, I expect to see clear differences between miRNA families that target the same protein or proteins in the same pathway. Towards that end, a part of my work at UCLA has been to design an assay to measure the changes and patterns in pri-miRNA processing for every human pri-miRNA in parallel. I have taken some foundational steps in designing this method.

Results

In collaboration with Dr. Sri Kosuri and his student Taylor ward, we designed a live-cell reporter that could be integrated into the genome of “landing pad” HEK 293-T cells. The landing pad is a genetically engineered site that contains a CAGGS promoter and a BXB1 integrase site in the AAVS1 locus (Fig.1). The AAVS1 locus is knowns as a “safe-harbor”, where insertions

here do not appear to have negative impacts on the cell, and the region remains transcriptionally active.

Proper integration of the new live-cell reporter will activate transcription of eGFP and mCherry (Fig. 1). A pri-miRNA hairpin is in the 3'UTR of the *mCherry* expression cassette, therefore like the transfection-based assay, a decrease in *mCherry* indicates and increase in pri-miRNA processing. Selection with puromycin was performed for sixteen passages to ensure that the non-integrated reporter was no longer present.

In direct collaboration with Taylor Ward, we produced two cell lines with integrated reporters for pri-miR-9-1, and a processing-deficient reporter called pri-miR-9-1-M1 (pri-miR-9-1 –dead). Using flow cytometry, one can measure the fluorescent intensity of both fluorescent proteins and compare the two cell populations by their geometric mean (gm), or measure of central tendency (Fig. 2A). By plotting eGFP vs. eGFP/mCherry, one can easily compare the two reporters processing efficiency (Fig. 2A). High eGFP/mCherry will indicate a decrease in mCherry, and therefore an increase in pri-miRNA processing. We saw good separation between cell populations expressing the pri-miR-9-1 and the pri-miR-9-1-dead reporter (Fig. 2B), with a nearly 3.5-fold difference in the gm between the two populations.

We wanted to ask questions by measure processing of all pri-miRNA hairpins in parallel, therefore we made a library that included reporters for every annotated human pri-miRNA. The library includes two negative control hairpins for each pri-miRNA that have mutations at the Drosha cleavage site from the 5' and 3' side. Grant Shoffner prepared the sequences for the library, and Taylor Ward cloned the library. I transfected the library into HEK 293T cells and performed puromycin selection for 16 passages until no signal from non-integrated reporter remained. The idea is that each cell will contain the reporter for a different pri-miRNA

Next we wanted to see if we could see evidence of a heirarchy in pri-miRNA processing by identifying highly processed pri-miRNA hairpins from pri-miRNAs with low processing activity. With the help of the great staff at the UCLA Broad Stem Cell Center FACS Core, I sorted the high processing cells from the low processing cells for two round of selection. I saw that I could successfully bias the cell population towards high-processing vs low-processing hairpins (Fig. 3). We are in the process of optimizing a Next-Generations sequencing pipeline for this assay.

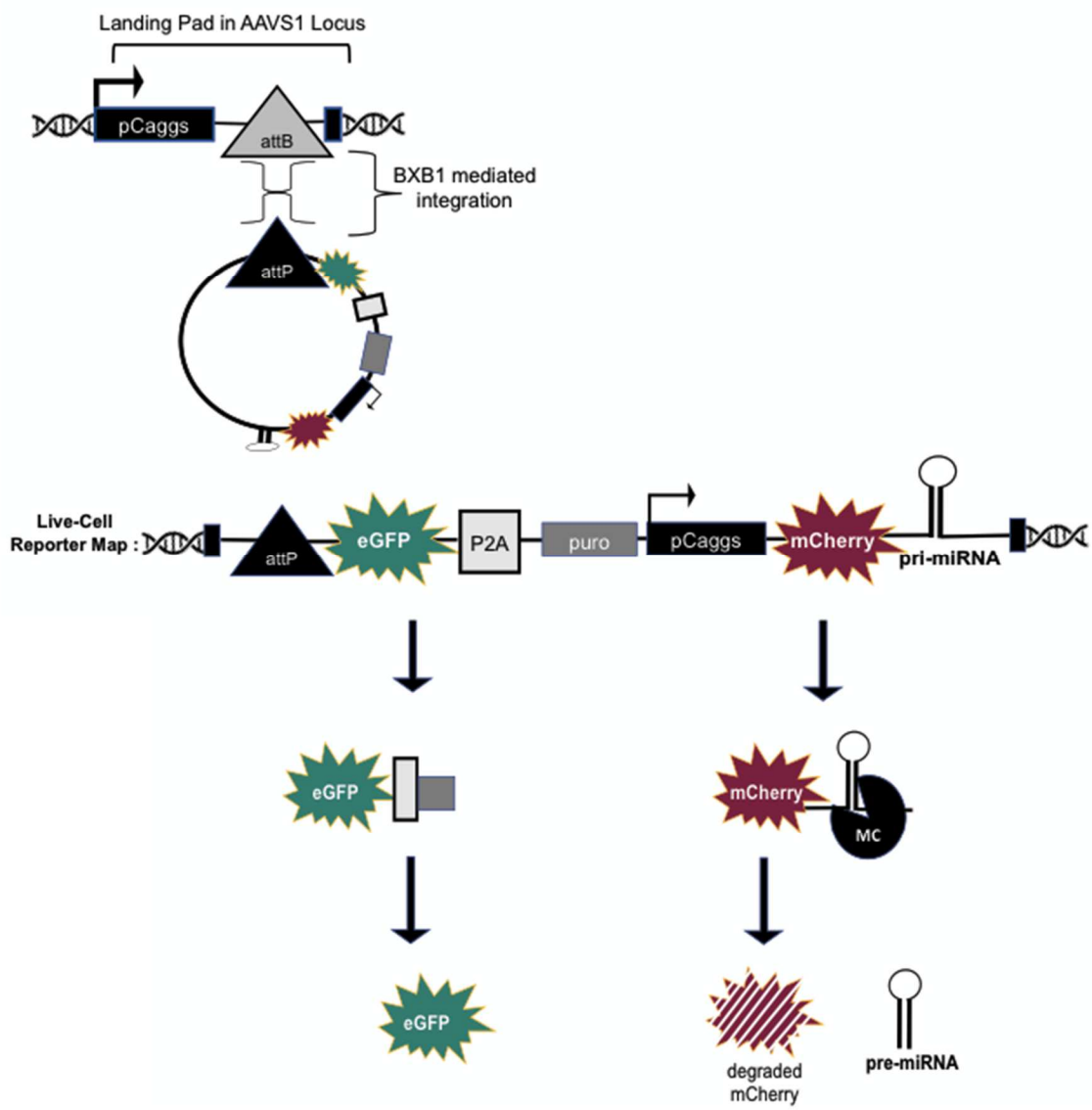


Figure 1. Live-cell reporter for integration into the AAVS locus.

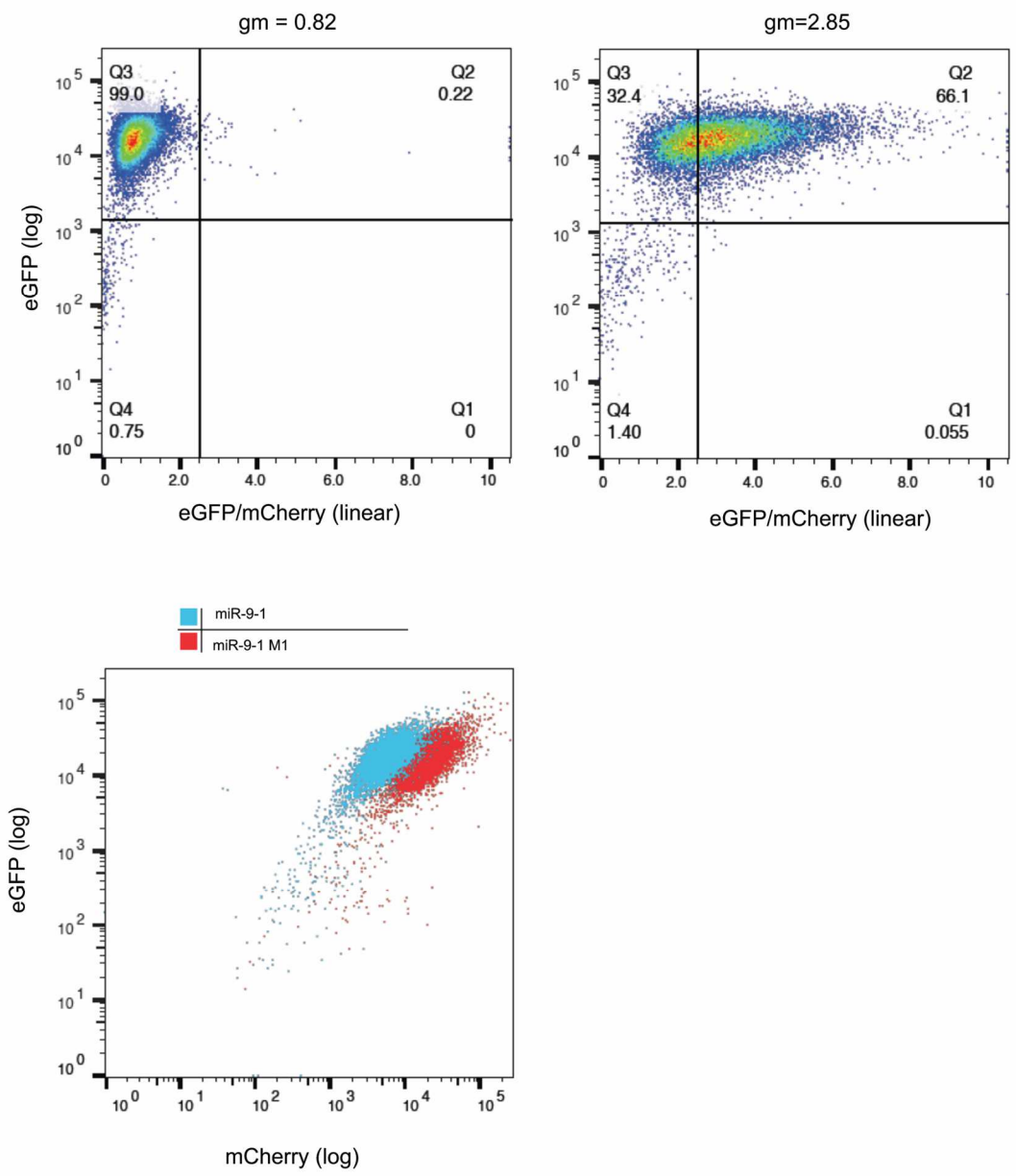


Figure 2. Flow-cytometry of the pri-miR-9-1 and pri-miR-9-1 dead cell lines. A-B) 30,000 cells are analyzed on eGFP and mCherry channels (gm=geometric mean). C) Comparison of the eGFP vs. mCherry intensity for both populations shows there is clear separation driven by differential processing of the two reporters.

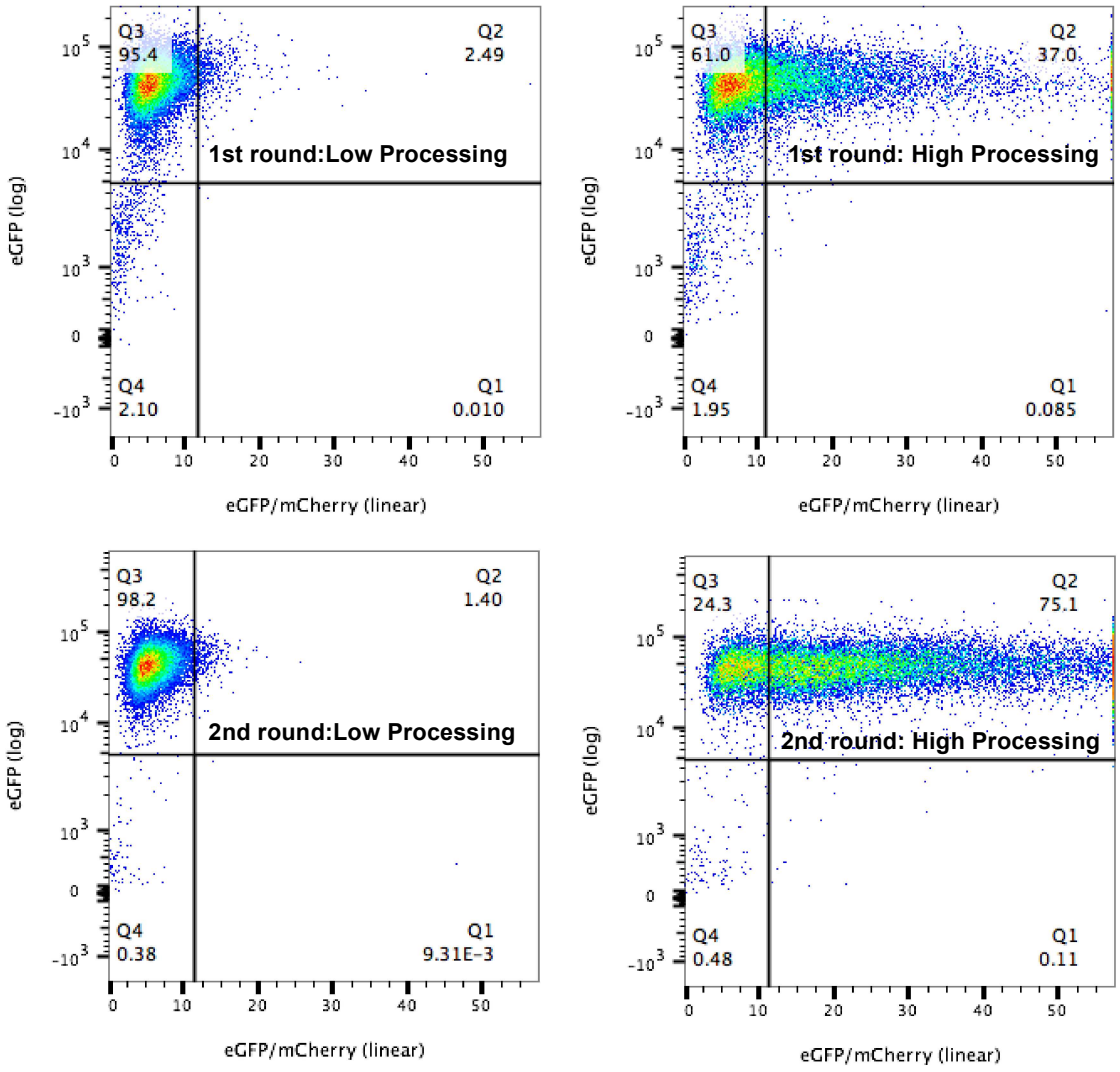


Figure 3. Cell sorting and selection of high-processing vs. low-processing pri-miRNAs

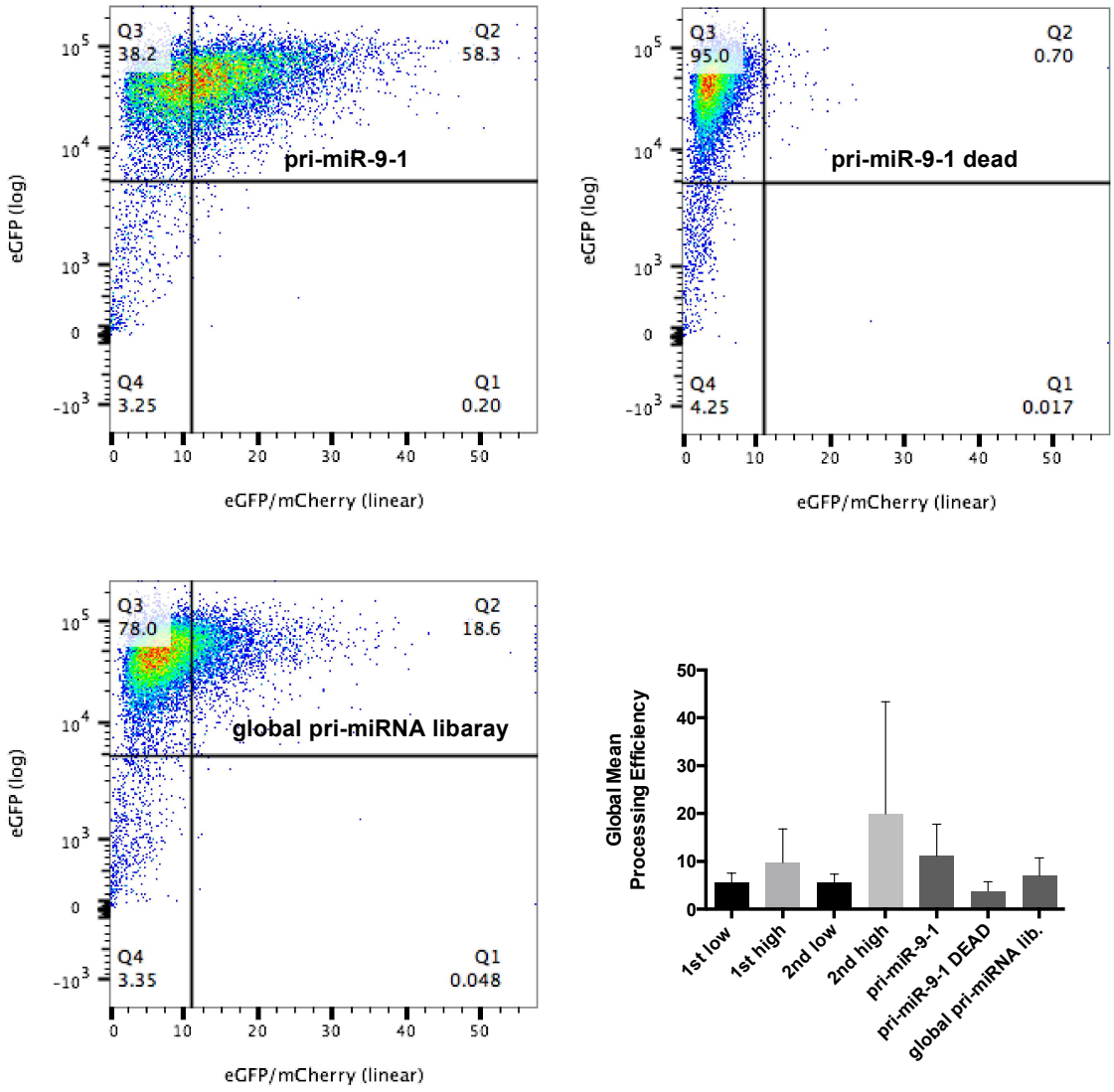


Figure 4. The Global pri-miRNA library can undergo selection

Procedure: flow pri-miRNA-processing assay

Program: BD FACs DIVA

<u>Experiment Set-up:</u>

1. Go to the “experiment tab in the browser and select “New Experiment”.

Use a blank experiment and name it over in the right panel.

2. To configure your settings, add a specimen by clicking the syringe symbol on the left of the browser. Add a tube to the specimen and name it “no color”
3. Make sure the tube is selected - it should have a green circle to the left of it.
4. Go to the parameters tab in the middle panel of the browser. Delete all the options but four. Change the remaining four configuration settings to “FSC” “SSC” “GFP” and “mCherry”. If the reporter contains other FP you will modify the two fluorescent channels to suit your FPs.
5. Unclick the lcg scale setting next to the two FP parameters – this puts these channels in linear mode
6. Go to the “Ratio” tab and add GFP as the numerator and mCHerry (or the FP with the hairpin in the 3’ UTR) as the denominator
7. Right click “Cytometer Settings” under the experiment you created. Select “Application Settings” and then, “Create worksheet”
8. Prepare the no color control by vortexing the filtered cells with the cap on for 2-3 seconds and load onto the cytometer
9. Press aquire and adjust the FSC and SSC voltages to place your population in the center of the graph. Make you correct for cell autofluorescence by adjusting the mCherry and eGFP voltages to be below the plus sign on the graphs
10. Once you are satisfied with your voltages for the “no color” control you can run the either the two-color control or the single color controls. Make sure the fluorescent signal is high and your voltages that you set for the FP channels are going to work. They should not need much adjustment if everything is going well. Linear range of the instrument is 0 to 10^5

11. Once your voltages are all set, right click on “Cytometer Settings” and select “Application Settings”. Save your settings with a keyword – not name or date (Felicia’s rules).

Compensations correction for the FP channels:

12. Go to the “Experiment” tab at the top of the browser and select “Compensation Set-up”. Then click “Create Compensation Controls”.
13. Click “ok” on the dialog box when it pops up
14. Expand the menu under “Compensation Controls” on the right, under your experiment folder
15. Place the pointer at the first “Unstained” sample (unstained is the same a no color)
16. Vortex and run the “no color” control again. Aquire and record 5000 events.
17. Run the “single color controls. While in aquire mode, adjust the P2 gates to surround the FP positive population of highy expressed mCherry or eGFP.
18. When gate is set, record 5000 events.
19. Select “ Experiment” at the top of the browser then “ Compensation Set-up” and then “Calculate Compensation”. If you have done everything correctly a box will pop up stating the calculation was successful. Mare sure that “enable compensation” is selected under the compensation tab in the middle of the browser.
20. Select “Apply Only” in the dialog box that pops up. You don’t need to save your compensation calculation because you will do it with each experiment

Collecting Data:

21. Create a new specimen again by clicking the syringe symbol at the top right of the browser. Specimen = general name of the sample set; Tube= specific sample name.

22. Toggle to a global worksheet by clicking the leftmost button on top of the worksheet browser (right screen).
23. Create the following plots in the global worksheet: SSC vs FSC scatter plot, GFP vs mcherry scatter plot (axis= log,log), eGFP histogram (biex), mCherry histogram (biex), ratio histogram (biex), GFP vs Ratio density or contour plot (log, biex).
24. Aquire and record 30,000 events for each sample. The parent population should contain at least 60% of your cells – get to know your population
25. You can print a pdf report of your raw plots by right clicking your experiment folder and selecting “Batch Analysis”.
26. Save FCS 3.0 files onto your USB for FlowJo analysis. You can chose to save your experiment as well. Felicia likes you to save your experiment on the computer as a backup.

FlowJo Analysis:

27. You can rent the FlowJo license usb from the FACS facility – if you rent from them you need to do your analysis there- can’t take usb out of the facility. Download the software and plug in the drive to activate
28. Drag the FCS files to the Flow Joe workspace
29. Highlight all files and click the “Derive Parameters” option under tools. A warning will pop up asking if you want to open a box for each FSC file – select ok.
30. Under the equation enter the equation = GFP/mCherry using the parameters button to enter the equations in the proper Flow Joe syntax. Ctrl+copy this equation to each equation box applying the ratio without scaling to each FSC file.
31. Open an empty layout to put all your graphs on

32. For each FSC file, make a SSC vs FSC plot with pseudocolor dot plots. Drop a gate around the parent population (P1) making sure to ignore cells with small FSC and large SSC scatter as these are most likely dead or dying cells. A good-looking plate of cells with have 60% or more falling in the P1 gate. Start your analysis with the negative control plate (dead hairpins).
33. Double click the cells in the P1 gate to analyze the P1 population
34. Create a 10% contour plot showing outliers of Derived (linear) vs.eGFP(log)
35. Drop a quadrant gate over such that the negative control population is isolated in Q3. Note the X,Y coordinants of the quadrant gate so you can apply it to all susbsequent files.
36. Copy and past the SSC vs FSC and the Derived vs Egfp to the workspace layout to batch export graphs

Appendix C: **Lessons from structural and biochemical studies on DGCR8**

I investigated many other questions about DGCR8 but did not obtained clear answers. My results may be useful for a future researcher and therefore are provided here. I also view this as an opportunity for academic reflection. One of my primary goals when I joined the Guo lab was to get a high-resolution 3D electron microscopy (EM) structure of DGCR8-pri-miRNA complex. This failed, but I did gather some preliminary data and I screened a number of mutations that could be of interest to future work. Another goal was to determine the precise DGCR8-contacting site on a pri-miRNA using an RNA footprinting approach. This approach has not succeeded, in part, because of the presence of heme in the DGCR8 protein. The heme-related interaction with footprinting reagents is interesting and worth discussion. Probably the biggest setback during my time at UCLA was my investigation into the redox properties of DGCR8. This work may prove instructive as it highlights a basic principle of sound experimental design: having good controls to cover all aspects of the conclusion. This is intended, mainly as a resource and therefore detailed protocols are included.

Footprinting analysis of the Rhed –pri-miRNA complex

Collecting structural information about the Rhed has been a major challenge that our group has been dealing with over the last decade. The domain has been recalcitrant to crystallization efforts and so we hoped to understand the Rhed-RNA binding surface using biochemical footprinting assays. Footprinting methods can be used to determine the binding site of a protein on an RNA. Proteins bound to an RNA protects the RNA from cleavage by chemicals or nucleases. I attempted to use hydroxyl radical, peryoxonitrite, RNAse, and x-ray footprinting approaches. Mainly the RNAse footprinting andt the x-ray foot printing were the most informative approaches. Here I present the higher quality data and summarize the results.

Results

Our goal was to identify the nucleotides that interact with the Rhed. Here I tested miR-23a-D, an apical junction model with and without the Rhed. I used RNase T1, which cleaves ssRNA after guanines, and RNAse V1, which non-specifically cleaves dsRNA. The cleavage pattern is very clear but there is no detectable protection. This is likely because the concentration of Rhed (0.2 -0.8 uM) is not high enough. The Rhed has considerably lower affinity for miR-23a-D (K_d =191 nM). When the complex is digested, there appears to be an increase in cleavage at G_{16} and G_{29} (Fig. 1**,** compare lanes 8,9 &13 to lane 7). This might indicate these nucleotides become more accessible to the T1 enzyme when bound by the Rhed. It is challenging to make a conclusion when the nucleotide positions cannot be clearly assigned. Technically, a longer gel would be preferred so cleavage across the entire RNA can be visualized.

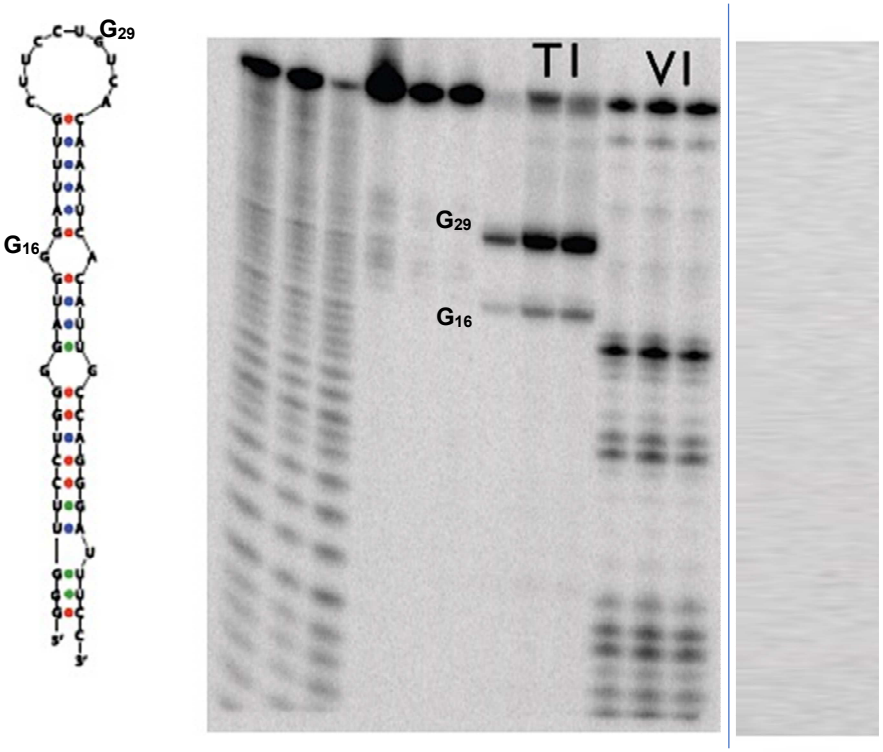


Figure 1: RNase footprinting of Rhed-pri-miR-23aD complex. On left: secondary structure prediction of pri-miR-23aD**.** On right: 12% denaturing gel of footprinting reactions. * indicates RNA has gone through the same treatment as the complex footprinting reactions. Experiment found on page 145 and 151 of labnotes #7

I tried to improve upon the RNAse footprinting by using the Rhed-pri-miR-21 complex. The Rhed binds pri-miR-21 with the highest affinity of all the pri-miRNAs tested K_d= 50nM (Quick-Cleveland). I wanted to be able to assign each position in the RNA, so I included a T1 and U2 RNase ladder. RNAse U2 cleaves preferentially cleaves purines. The ladders helped greatly with assigning the nucleotides of pri-miR-21. As before, no clear protein footprint was observed. Looking at the T1 cleavage of the Rhed-pri-miR-21 complex (Fig. 2 compare lanes 8 & 9 with lanes 10 &11) it appears again that the protein is enhancing the cleavage of certain nucleotides (marked with pink star). The V1 cleavage of the complex echoes this pattern, showing enhanced cleavages in the presence of the Rhed (Fig. 2, compare lanes 12 & 14 with lanes 13 & 15). There are four main points of enhanced cleavage that I feel confident in assigning; A_{53}, $G_{81,}$ $G_{91,}$ and, $C_{104.}$ The data shows that these nucleotides become more senstitive to RNAase cleavage in the presence of protein.

We attempted x-ray foot-printing of the NC1-pri-miR-30a complex in collaboration with Dr. Corie Ralston and Dr. Sayan Gupta from Lawrence Berkeley National Laboratory. This technique provides high-resolution residue-specific information about structure and dynamic by generating hydroxyl radicals *in situ* (1,2). This approach seemed ideal because it did not require the addition of enzymes or chemicals. For this experiment, I used a pri-miR-30a construct that has a primer-binding site at the 3' end (pri-miR-30a-SHAPE). To visualize the cleavage products, I used an end-labeled primer (p.681) and performed reverse transcription (RT). Cleaved positions in pri-miR-30a-SHAPE are indicated by truncated RT-transcripts. These reactions were run on a 20% sequencing gel. The results seem to suggest that NC1

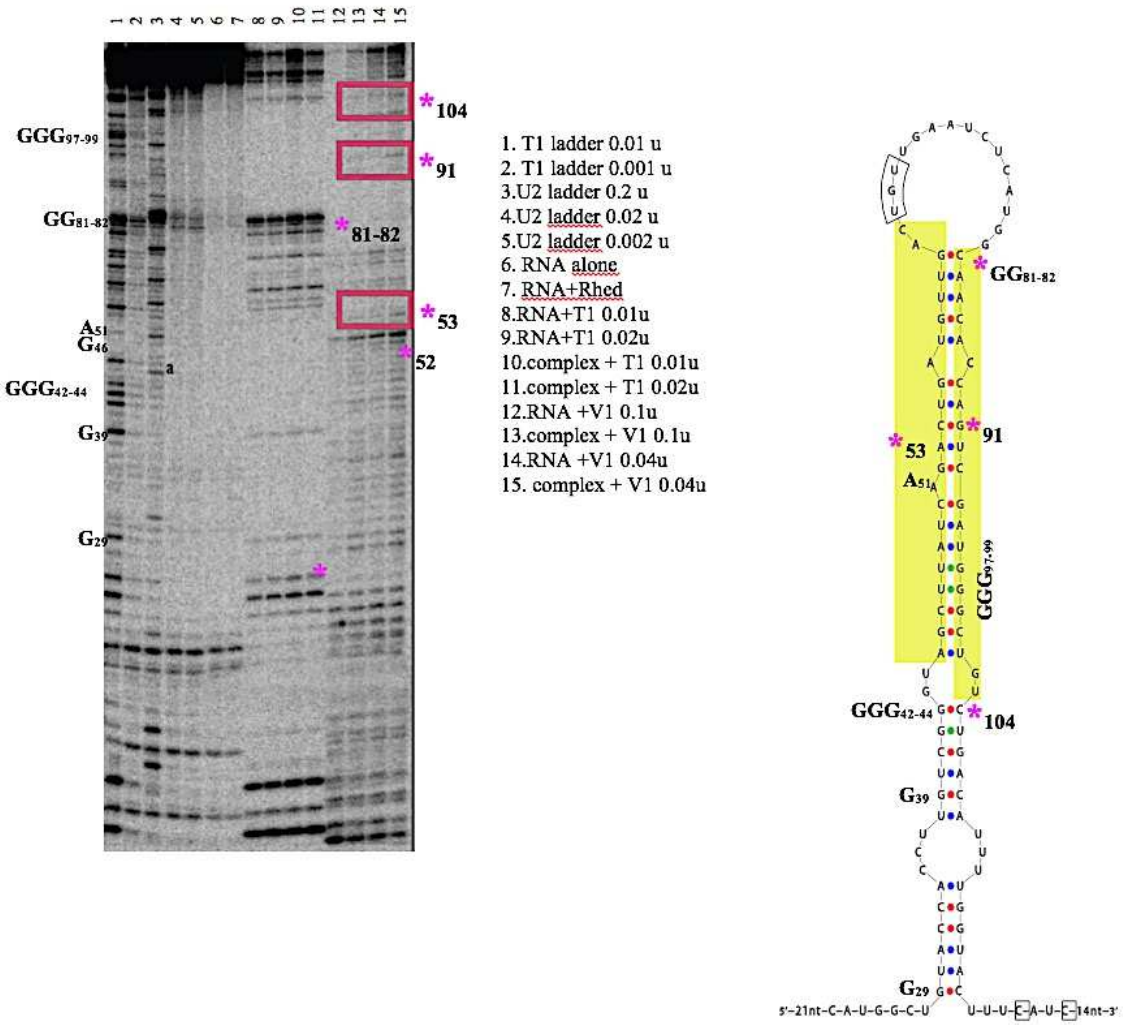


Figure 2. RNase footprinting for the Rhed-pri-miR-21 complex. On left: 20% denaturing gel of footprinting reactions. Assigned resides based on T1 and U2 digestion ladders are labeled. Pink stars mark residues show differences in cleave in the presence of the Rhed. On right: secondary structure prediction of pri-miR-21L1 with assigned residues labeled. Experiment on pages 4-6 of labnotes #7. protects pri-miR-30a-SHAPE extensively (compare lanes 1-4 to 5-8). There appear to be a dose-dependent increase in cleavage as exposure time to the x-ray increases (compare lanes 1-4). As the RT proceeds into the hairpin region the resolution goes down. It is hard to draw a conclusion from this experiment. This is the best x-ray foot-printing gel I produced.

I synthesized peroxynitrite by reacting sodium nitrite with hydrogen peroxide, cleaning up residual peroxided with the addition of manganese oxide. RNA cleavage by peroxynitrite looked promising in that I did see clean cleavage patterns (data not shown, page126 labnotes #6 for best gel)(3). After many attempts of footprinting, I incubated the mcherry-tagged zebrafish Rhed with some of my synthesized peroxynitrite to see if it was stable in the footprinting condition (Fig.4). Immediately upon addition to the protein, I observed changes in the electronic absorption spectra. These changes were easiest to see at 587nm (the mCherry tag) because the 450 nm söret peak is under a large absorbance from peroxynitrite. This led me to suspect that peroxynitrite might be interfering with the Rhed.

We entertained the idea of doing phoshporothioate foot-printing where one can add phosphothioated uracil to the transcription reaction. Iodine is then added as a footprinting reagent to probe for phosphate groups involved in RNA-protein interactions. My first step was to test how NC1 tolerated addition of iodine because of the potential for oxidation of cysteine residuesto cysteic acid (iodination). I added 1mM iodine to NC1 and observed no change in the electronic absoption spectra over 3 hours. I increased this to 3mM and observed a change in the 450 nM söret peak of NC1 (Fig. 5). This could not be rescued with addition of DTT. From this we decided it prudent not to pursue this method further.

Summary

My attempts to map the DGCR8-pri-miRNA interaction using a footprint approach was most successful using RNase protection assays and x-ray footprinting. Additionally I tried hydroxyl radical, peroxynitirite, and phosphorothioate probing of RNA-Protein interactions. Peroxynitrite and iodine (used in phosphorothioate footprinting) potentially interfere with DGCR8 heme-binding. I did not observe a footprint for the Rhed-pri-miRNA complex. It may be that the

pri-miRNA remains highly accessible when bound to the Rhed. Alternatively, because the Rhed-pri-miRNA interaction is lower affinity my assay conditions may not have been ideal for complex formation. NC1 has a higher affinity for pri-miRNAs and seemed to produce a very large footprint spanning most of pri-miR-30a-SHAPE. If further x-ray footprinting is pursued for this complex, I recommend forming the complex with end-labelled pri-miRNA to avoid the reverse transcription step. Technical challenges stalled this project. I optimized RNA labelling and RNA-precipitations protocols during the course of this work. I hope they might be useful in the future.

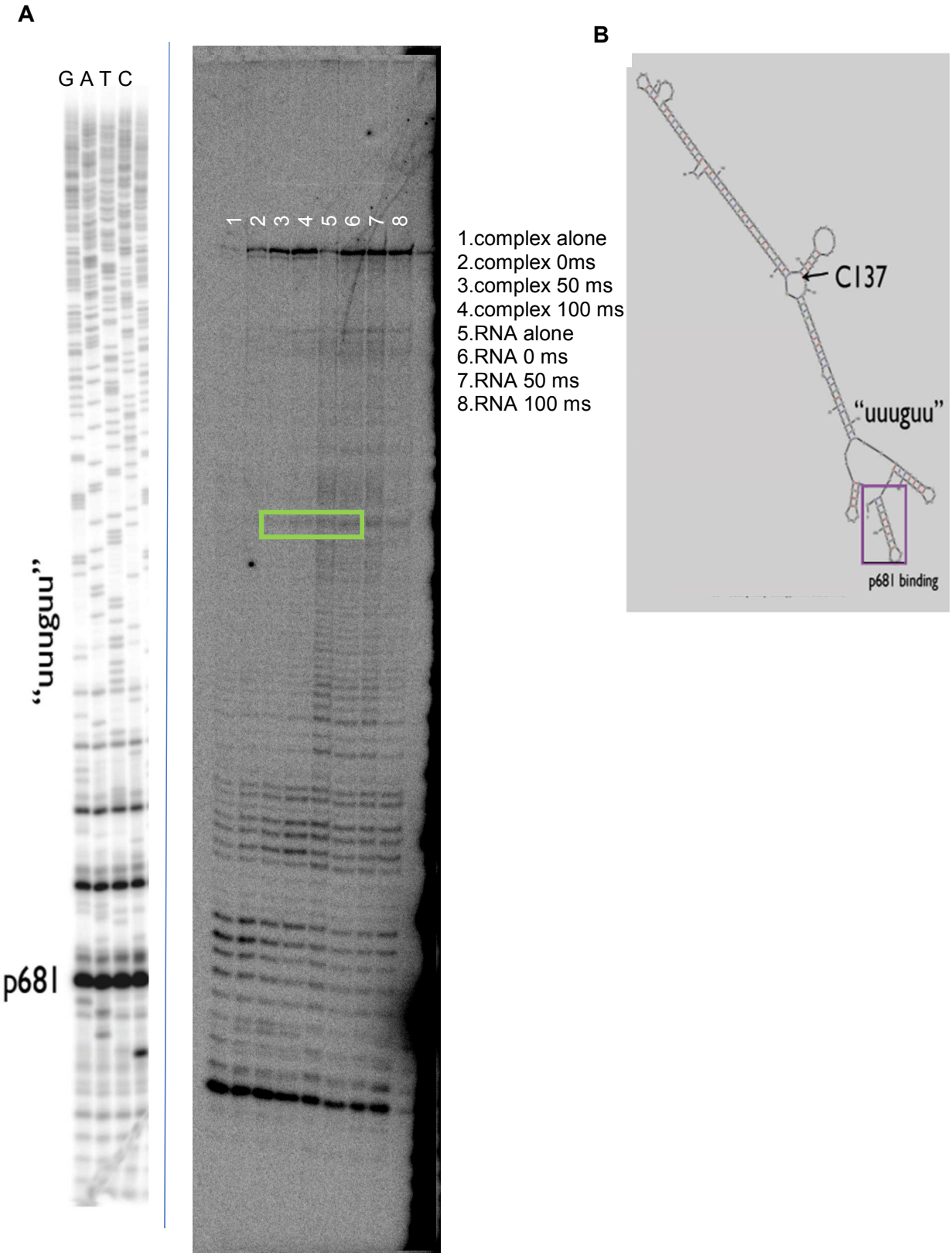


Figure 3. X-ray footprinting of the NC1-pri-miR-30a-SHAPE complex. A) 20% sequencing gel with sequencing ladder and x-ray footprinted samples.C_{137} is highlighted in green.

B) Secondary structure prediction of pri-miR-30a-SHAPE showing the position of primer annealing. The "uuuguu" is labeled on both gel and RNA secondary structure for orientation. Experiment on page 87 labnotes #6 for complex preparation. Experiment on page 3 labnotes #7 from sequencing ladder preparation.

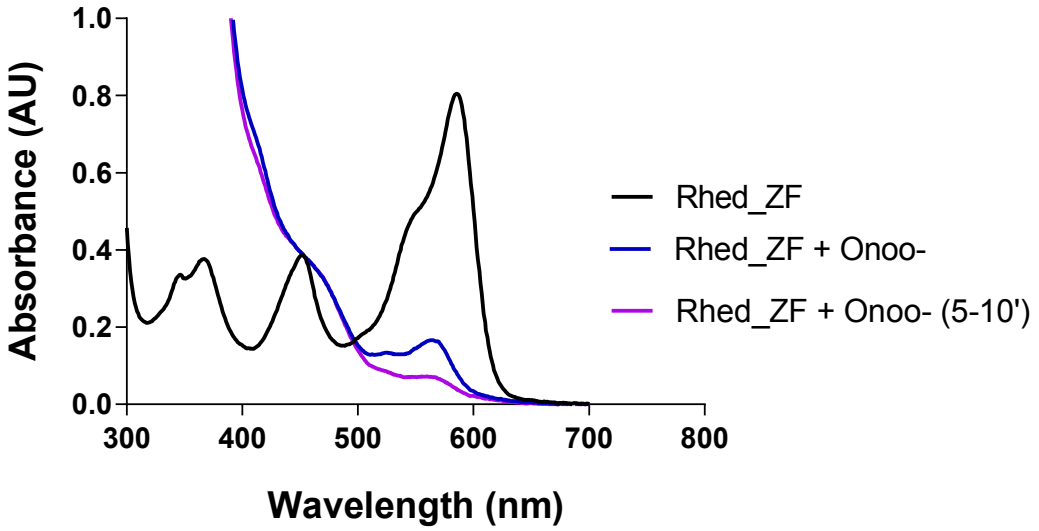


Figure 4. Incubation of the zebrafish homologue of the Rhed with peroxynitrite. The mCherry-tagged zebrafish Rhed was incubated with synthesized peroxynitrite anion (Onoo-) and the electronic absorption spectra was recorded immediately after addition, as well as 5-10 minutes later.

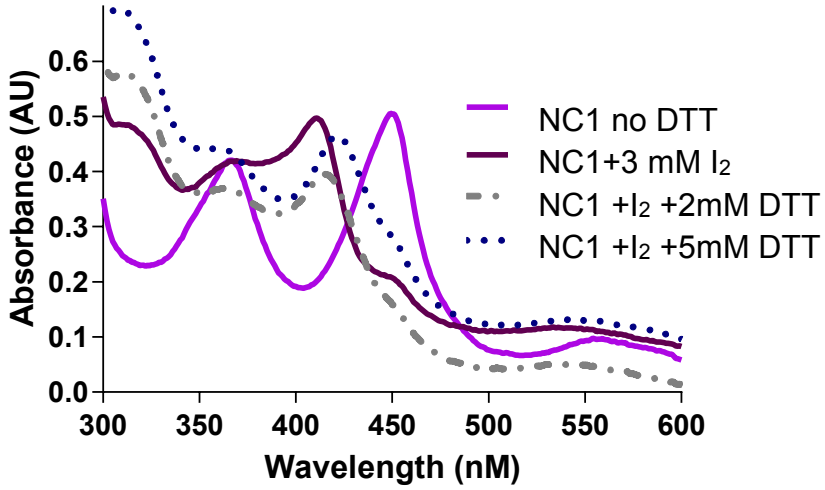


Figure 5. Incubation of NC1 with iodine. NC1 incubated with iodine. Addition of DTT does not reverse disruption of the double-cysteine heme ligation.

Cryo-EM of the NC1-pri-miRNA complex

Our group solved a low resolution three-dimensional of the DGCR8-pri-miR-30a complex using electron tomography (4). In a collaboration with Dr. Hong Zhou and his student, Ke Ding we hoped to achieve a high-resolution structure using cryo-EM. Here I present some preliminary images and two-dimensional class averages, and summarize the results.

Results

My colleague Grant Shoffner had done extensive work generating DGCR8 mutants that we hoped would be more amenable to structural studies. One such mutant was NC1-EE2, where Grant had engineered a FAB binding site in to the acidic loop of the Rhed. The NC1-EE2-pri-miRNA complex peak appeared to be sharper than usual, potentially indicating a more stable and homogenous complex (Fig. 4a). I looked at some complex samples by negative stain (1% uranyl-acetate), and the complex appeared mostly homogenous (Fig. 4b). At this point we began our collaboration with Ke Ding and Dr. Hong Zhou. Subsequent negative stain images appeared to be *positively* stained (Fig.6). This could indicate the stain was applied too thinly or that the stain interacts with the particles. Both situations are not ideal, however we planned to use negative stain as a tool to assess homogeneity and density and continue onto cryo-EM. The cryo-EM grids prepared by Ke Ding looked good enough for imaging in that the particle density was not too high or low, and the ice thickness appeared uniform (Fig. 7). However, after imaging on the Titan Krios and selecting particles for analysis we saw virtually featurless 2D class averages (Fig. 8). This was the strongest indicator we had that our NC1-ee2-pri-miRNA complex was not high enough quality to continue the study.

I tested two common EM stains, uranyl acetate and uranyl formate, to assess how they interact with the Rhed. First, I incubated 2 uM Rhed with ~100% heme occupancy with the standard working concentration of 1% uranyl acetate (UA). I observed dark brown precipitate in

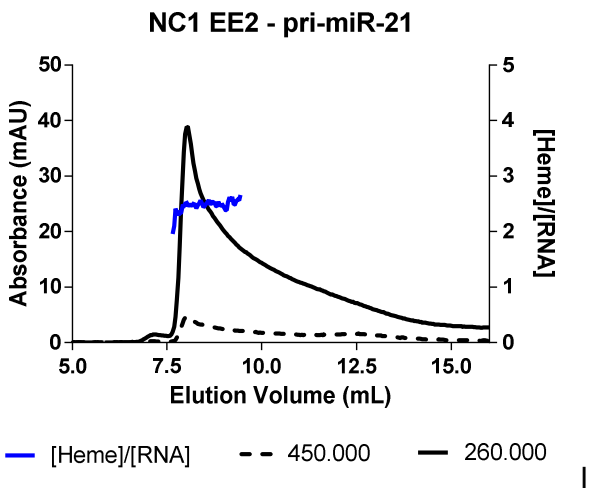
the cuvette and noted the absorbance from the stain obscured the 450 nm söret peak. I reduced the UA concentration to 0.025% and 0.075% uranyl acetate (Fig. 9A). The 0.025% UA condition still has a clear 450 nm peak but increasing the concentration to 0.075% UA swallows the 450 nm peak. Both conditions caused precipitation of the protein. I estimated pH of the UA stain using pH strips and it was very low at ~pH 4. The Rhed doesn't do well in below pH 6.0 (Fig. 10), however the buffer should have protected the protein from a large shift in pH. From this I concluded that the UA stain is not optimal for study of the Rhed.

I then incubated 4uM Rhed and NC1 with 1% uranyl formate (UF). This stain is supposed to be a better negative stain for resolution of smaller complexes. This time I included 0.025% UF in baseline so it wouldn't obscure the protein peaks. This caused some problems with the baselines for Rhed+UF and NC1+UF (Fig. 11). I observed brown precipitate in both the Rhed+UF and NC1+UF conditions (data not shown). Intriguingly, when I form the NC1-pri-miR-21 and the Rhed-pri-miR-21 complex there was no visible precipitate in the cuvette. Additionally, the Rhed-pri-miR-21 complex electronic absorption spectra indicated no change in the söret peaks. This may indicate that although DGCR8 is sensitive to the EM stains, in complex with RNA the heme is somewhat protected.

Summary

Without another good DGCR8 construct to try, this project stalled. The Drosha-DGCR8-pri-miRNA complex would be a bigger structural target that might provide more features to align particles for better resolution. Cryo-EM seems like a viable approach that should be continued. If negative stain is used to assess complex quality, I recommend using the lowest as possible concentration of UF.

A



B

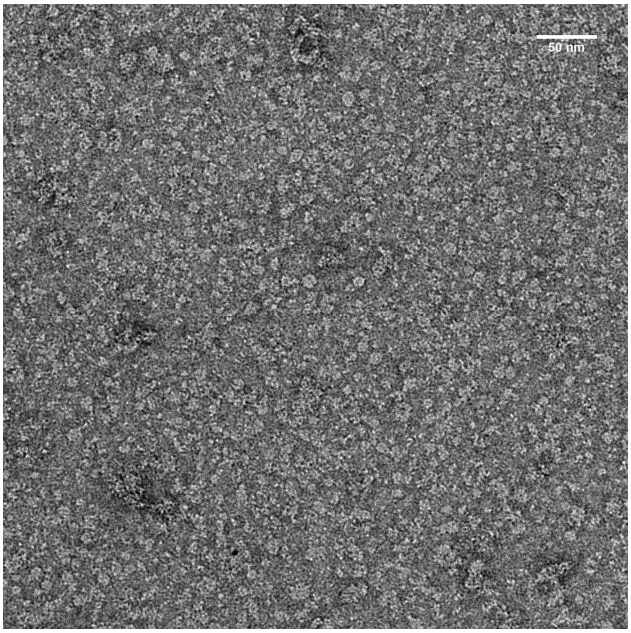


Figure 6. The NC1-EE2-pri-miR-21 complex. A) Complex purification by size-exclusion chromatography. B) Negative stain (1% UA) image of the complex, scale bar indicates 50 nm.

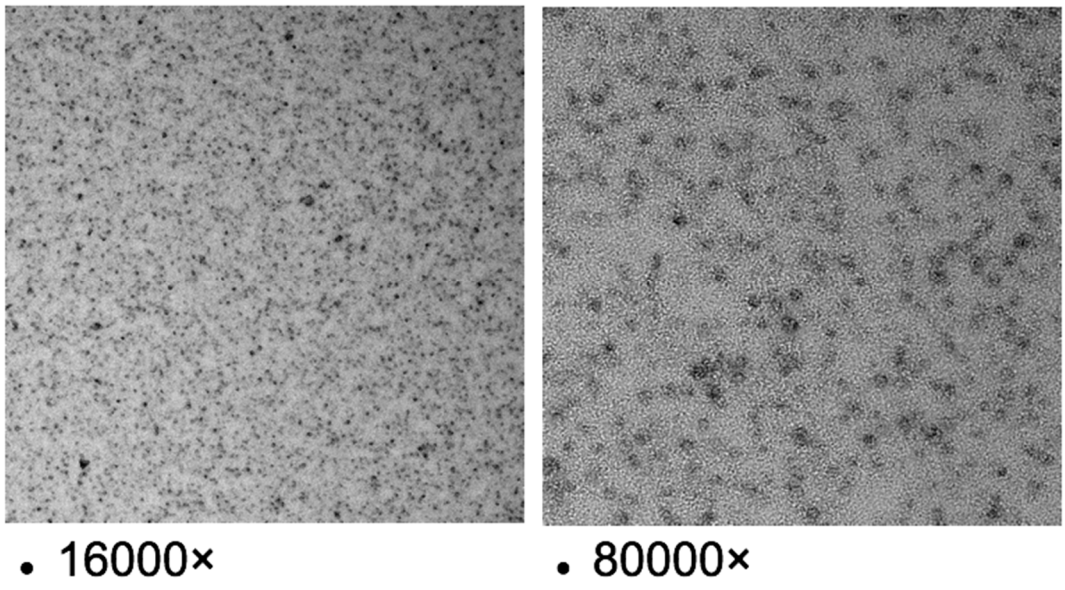


Figure 7. Negative stain (1% UA) images of the NC1-EE2-pri-miR-21 complex (Ke Ding).

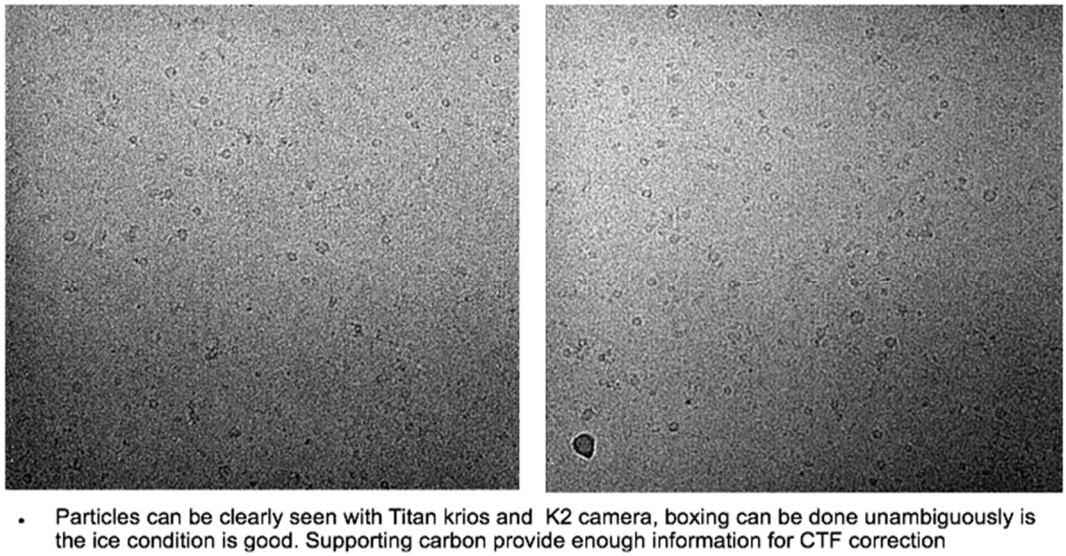
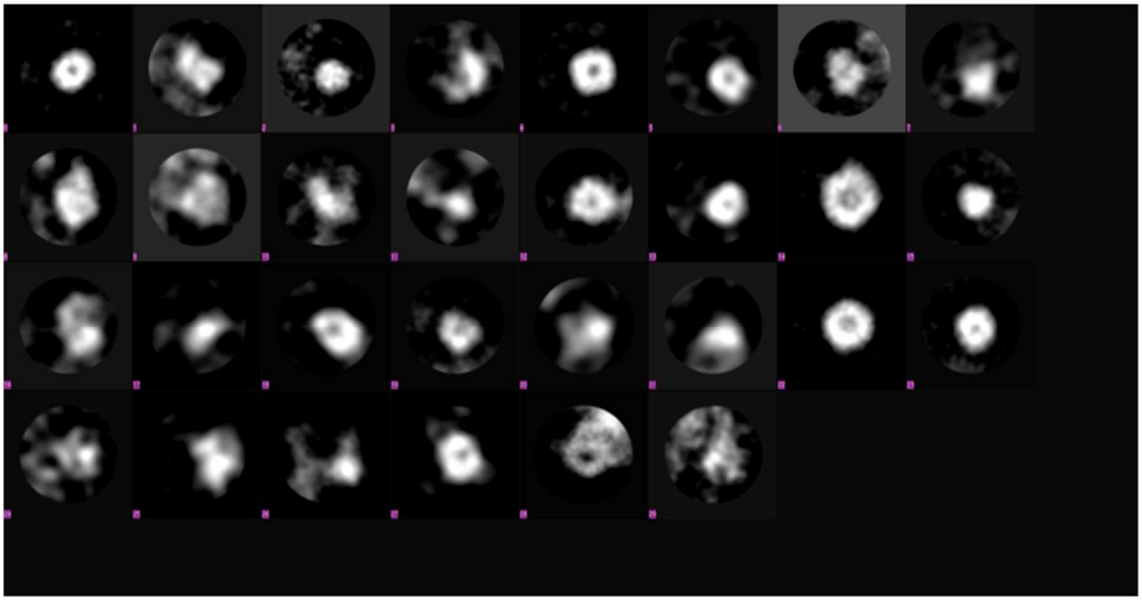
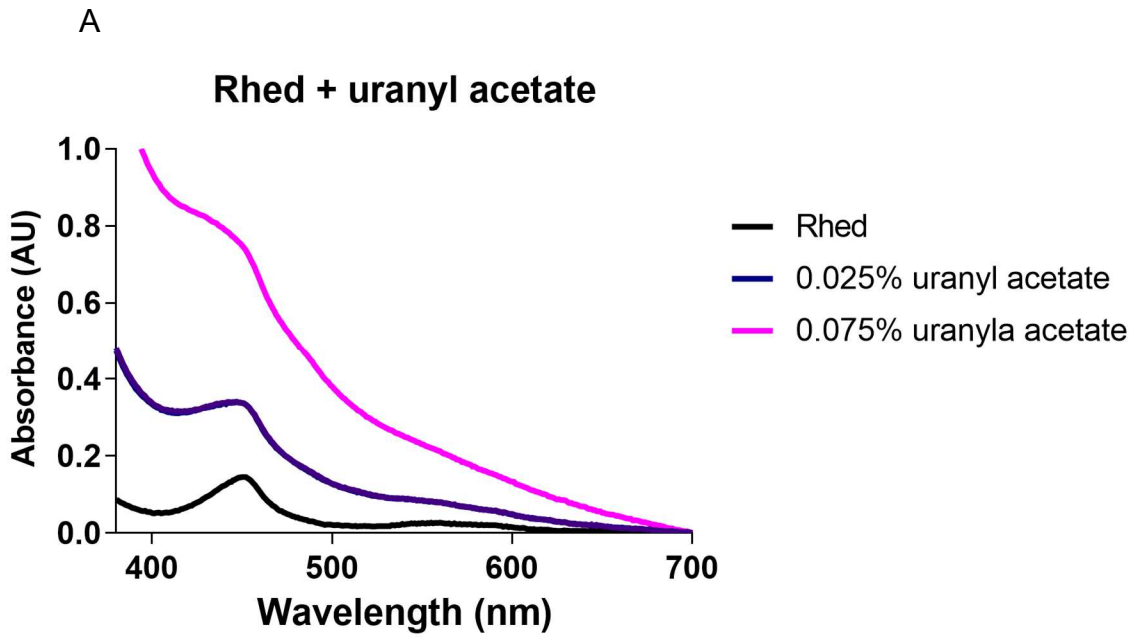


Figure 8. Images of cryo-EM grids with NCI-EE2-pri-miR-21 complex (Ke Ding).



- Features can be seen (donut), but images could not be properly aligned. There is no need to go to 3D if 2D class averages are bad for now.

Figure 9. 2d class averages of selected particles.



B

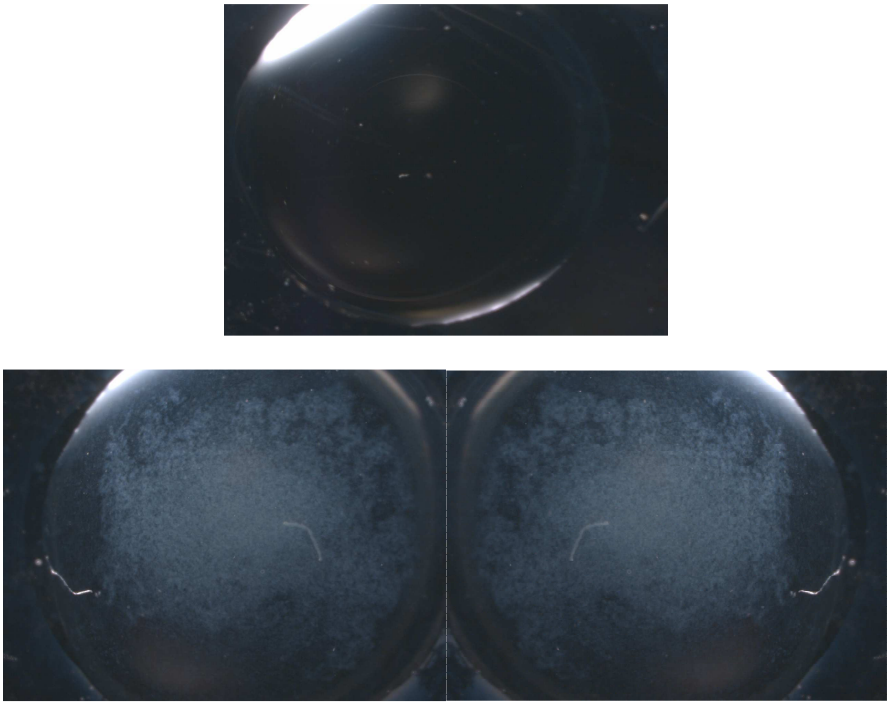


Figure 10. The Rhed interaction with UA. A) Electronic absorption spectra of the Rhed with 0.025% and 0.075% UA. B) 10x magnification picture of UA 0.025% solution alone (top right), Rhed with 0.025% UA (bottom right) and Rhed with 0.075% UA (bottom left).

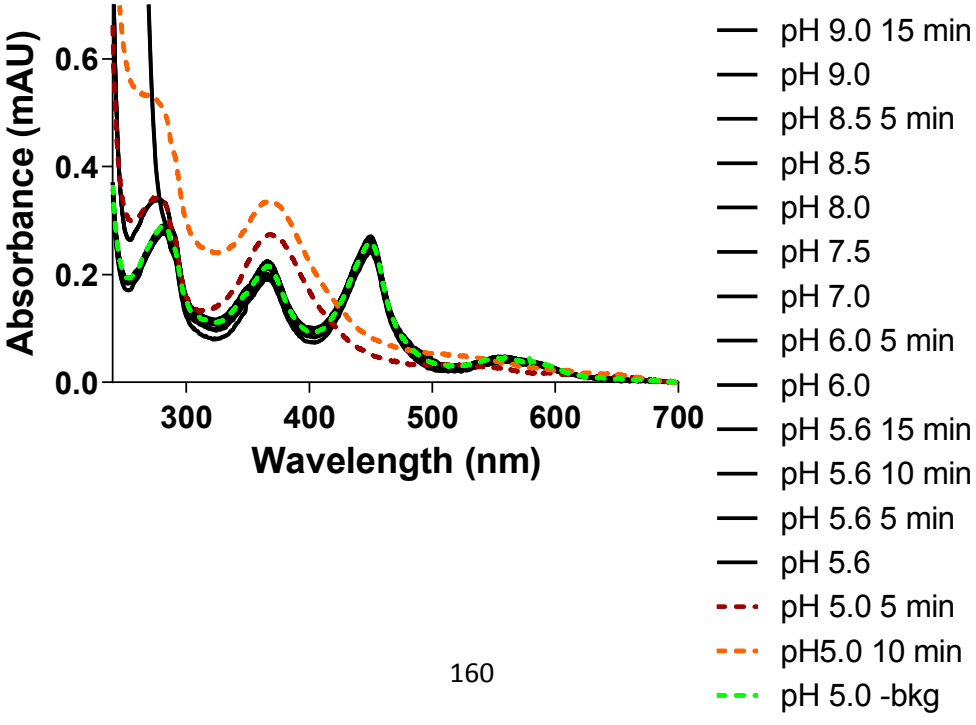


Figure 11. The Rhed in different pH conditions. Electronic absorption spectra of the Rhed incubated in various pH conditions. Buffers that had no visible effect on the Rhed absorption are black, while conditions that caused changes are colored.

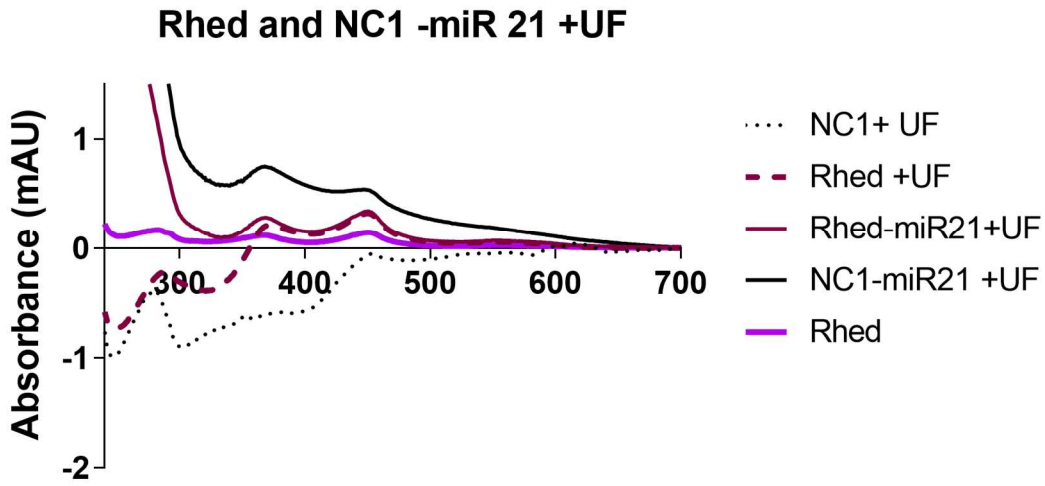
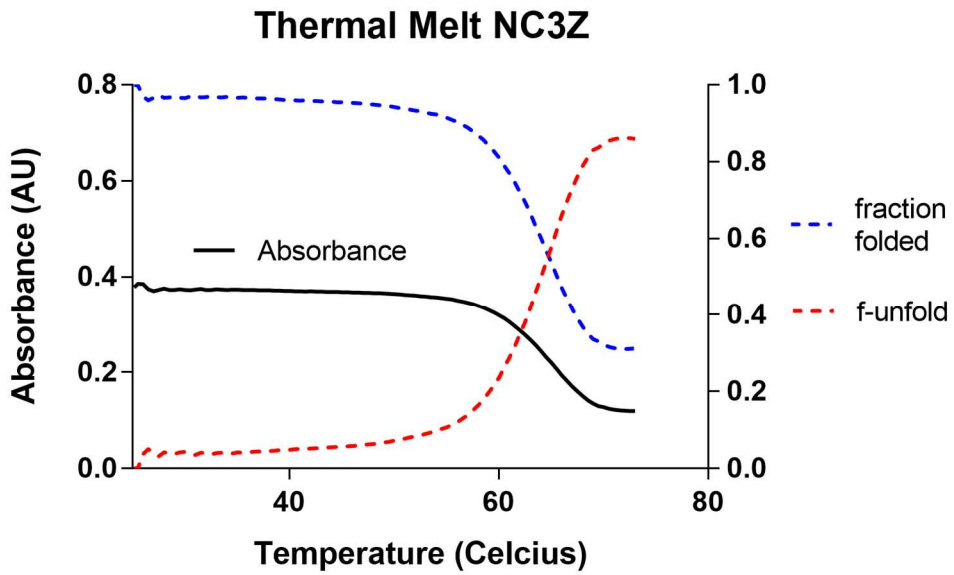


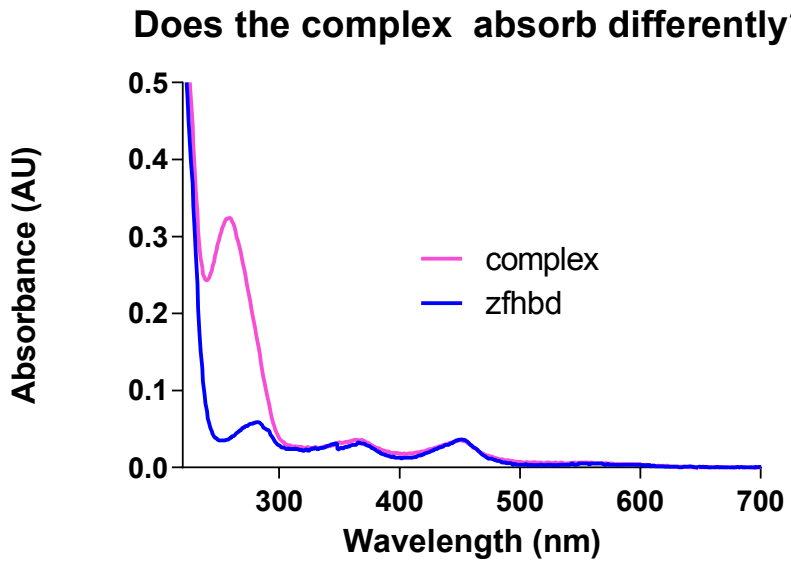
Figure 12. NC1 and Rhed with 0.025% UF. Electronic absorption spectra of NC1 and the Rhed with 0.025% UF and the NC1 and Rhed in complex with pri-miR-21 with 0.025% UF.

Other useful biochemistry experiments with the Rhed

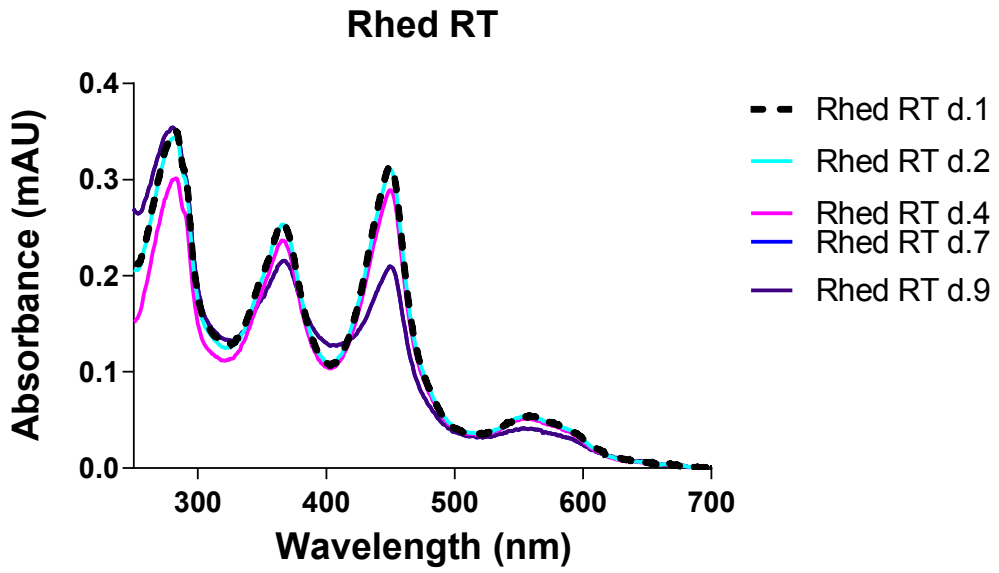
A



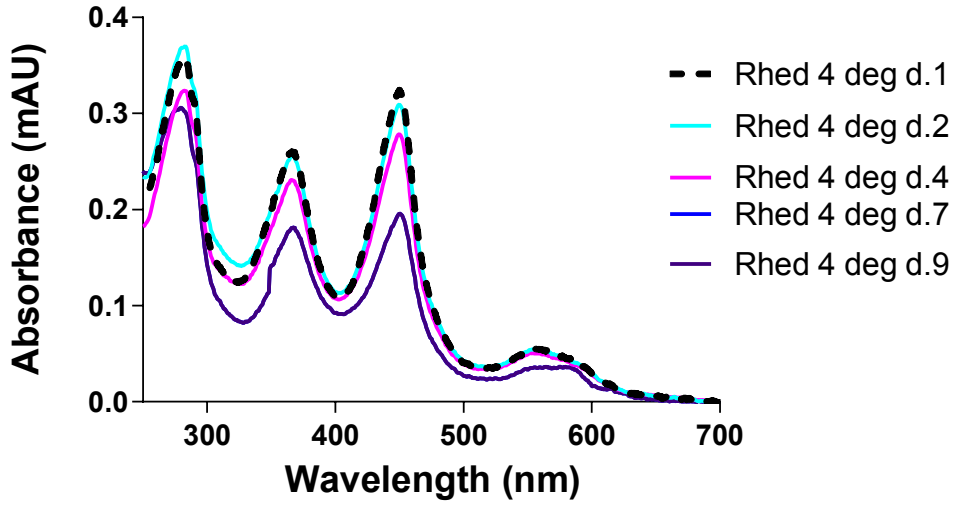
B



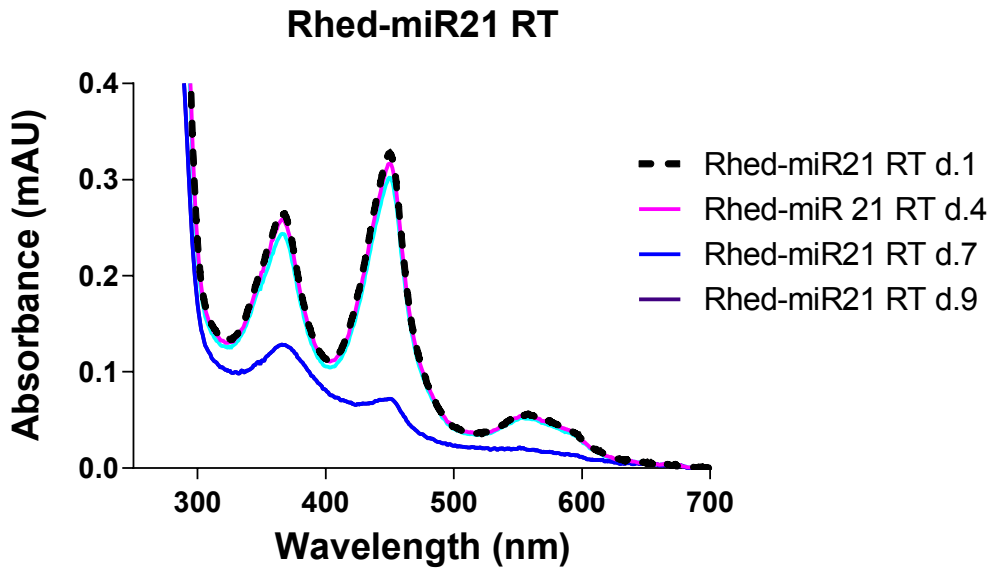
C


Rhed RT
Absorbance (mAU)
0.4
0.3
0.2
0.1
0.0
300
400
500
600
700
Wavelength (nm)
Rhed RT d.1
Rhed RT d.2
Rhed RT d.4
Rhed RT d.7
Rhed RT d.9

D


Absorbance (mAU)
0.4
0.3
0.2
0.1
0.0
300
400
500
600
700
Wavelength (nm)
Rhed 4 deg d.1
Rhed 4 deg d.2
Rhed 4 deg d.4
Rhed 4 deg d.7
Rhed 4 deg d.9

E


Rhed-miR21 RT
Absorbance (mAU)
0.4
0.3
0.2
0.1
0.0
300
400
500
600
700
Wavelength (nm)
Rhed-miR21 RT d.1
Rhed-miR 21 RT d.4
Rhed-miR21 RT d.7
Rhed-miR21 RT d.9

F

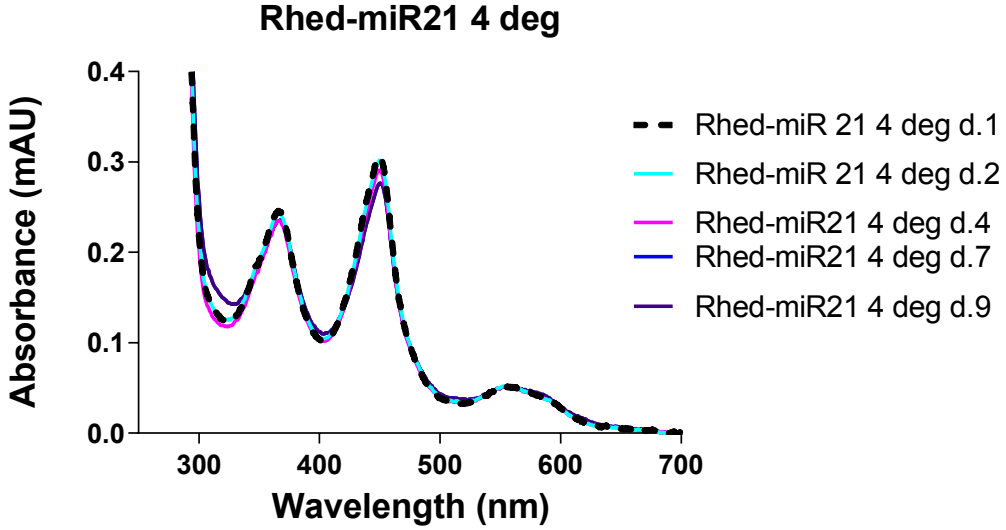

Rhed-miR21 4 deg
Absorbance (mAU)
0.4
0.3
0.2
0.1
0.0
300
400
500
600
700
Wavelength (nm)
Rhed-miR 21 4 deg d.1
Rhed-miR 21 4 deg d.2
Rhed-miR21 4 deg d.4
Rhed-miR21 4 deg d.7
Rhed-miR21 4 deg d.9

Figure 13. Biochemistry with the Rhed. A) Melting curve analysis on the CARY-300 thermal melt program shows the Rhed melts at 56°C. B) Electronic absorption spectra of the zebrafish Rhed and the zebrafish Rhed-pri-miR-21 complex. C) Analysis electronic absorption spectra of the Rhed stored at room temperature (RT) over 9 days. D) Analysis of the electronic absorption spectra of the Rhed stored at 4°C. E) Analysis of the electronic absorption spectra of the Rhed-pri-miR-21 complex stored at RT. F) Analysis of the electronic absorption spectra of the Rhed-pri-miR-21 complex stored at 4°C.

MC activity in low oxygen

As previously discussed, Drosha and DGCR8 are both heavily regulated post transcriptionally (Chapter 1 & 4). DGCR8, as a hemoprotein, also is regulated by heme-availability (Chapter 3). We were curious if we could identify changes pri-miRNA processing in response to physiological stimuli. Ever since the discovery of the specificity for Fe(III) heme, many in our group have been excited by the idea that the MC may respond to redox regulation. My colleague Rachel Senturia, did extensive work in this area, but the results were not reproducible. With the development of the live-cell assay (5), it became easier to ask question about the physiological response of the MC. I began testing many chemical stimuli related to redox. I tested hydrogen peroxide, nuclear-targeted hydroxyl radicals, 2-AAPA (glutathione reductase inhibitor), a series of electron transport chain inhibitors (rotenone, oligomycin, & antimycin) and hypoxia (1% oxygen). Exposure of cells to hypoxic conditions appeared to increase pri-miRNA processing. Here I present the data that led me down a rabbit hole that cost me nearly two years of data collection. The importance of having multiple controls to cover all alternative hypotheses cannot be overstated. Even now as I revisit this data, I still cannot explain the contradictions.

Results

Changes in pri-miRNA processing activity that we have measured are generally subtle. Therefore, we were surprised when 1% oxygen stimulated endogenous MC activity (Fig. 1A). Overexpression of DGCR8 showed even larger changes in slope, while overexpression of the C352A did not respond to changes in oxygen levels. This indicated that the increase in MC activity in response to 1% oxygen could be due to the presence of Fe (III) heme. However, DGCR8 protein expression went down in 1% oxygen (Fig. 1B). This downregulation of DGCR8 in low oxygen was also reported in breast cancer cells lines and in pulmonary artery smooth

muscle cells (6,7) Despite this, the mature miRNA levels from reporter increased, as shown by Northern blot analysis from two different cells lines using two reporters (Fig. 1C & 1D). I noticed some variation in this induction of MC activity in low oxygen. Counterintuitively, spikes in reactive oxygen species (ROS) production have been measured , and lower pO_2 can cause oxidative stress(8,9). Further, hypoxia induces HiF-1α as well as a ROS and N-acetyl cysteine (NAC) which acts as a ROS scavenger, inhibited ROS production as well as HiF-1α activation in MCF-7 cell (10).

I wondered whether a transient increase in ROS might explain the increase in MC activity and I predicted if that were true the increase in MC activity would be time-dependent. To test this idea, I performed a time course using the live-cell assay in 21% and 1% oxygen (Fig. 13). Each time point was a separate plate of cells with identical transfection reactions. The results showed MC activity was much more stable in 21% oxygen with only mild variations in eYFP vs mCherry slope (Fig. 13A). In 1% oxygen, there were larger fluctuations in MC activity, the peak being at ~16 hours post-induction and later ~48 hours (Fig. 13B). At all-time points, MC processing activity was greater in low oxygen. These results suggested to me that ROS might be involved in the enhanced MC activity. To explore this further, I measured MC activity in 21% and 1% oxygen in the presence of NAC and dithiothreitol (DTT), which are both expected to attenuate production of ROS (11). My results showed that in 21% oxygen, NAC and DTT had no effect on MC activity (Fig. 14). However, in 1% oxygen, I observed a dose-dependent decrease in MC activity with the 5 mM NAC + 0.5 mM DTT having the largest negative effect (Fig. 14). Cell in a 1% oxygen +0.5 mM DTT condition all died. This may be related to the link between DTT and ER-stress, but I did not continue experiments with DTT. At this point, I felt I had enough evidence to support ROS as a key player in this low oxygen-MC story.

A common way to induce a “hypoxic state” other than a low oxygen incubator is through chemical means. I measured the effects of $CoCl_2$, desferoxamine mesylate (DFO), and dimethyoxaloylglycine (DMOG) on MC activity using the live-cell assay. $CoCl_2$ treatment mimics

hypoxia by disrupting the interaction between HiF-1α and the von Hippel-Lindau protein that helps degrade HiF-1α in normal oxygen conditions (12). DFO is an iron chelator that inhibits prolyl-hydroxylases that normally repress the expression of HiF-1α (13), DMOG is a 2-oxoglutarate analogue that disrupts prolyl-hydroxylase activity (14). The experiments showed that treatment of cells with $CoCl_2$ and DFO resulted in a subtle dose-dependent increase in MC activity, while DMOG treatment had virtually no effect (Fig. 15). The interpretation of this experiment is not completely straight forward in that DGCR8 binds and is activated by Co(III) (15) , and chelating iron could have effects on DGCR8 as well. The DMOG experiment is probably the cleanest in terms of interpretation, and this suggests that the low oxygen induction of the MC is not dependent on HiF-1α.

I wondered whether this response to low oxygen could be related to DGCR8 heme-binding properties. To test this, I measured MC activity in 21% and 1% oxygen in heme-depleted media (HDM) and in HDM with the heme synthesis inhibitor, succinyl acetone (SA). I discovered that in both heme-depleted conditions the 1% oxygen stimulation of MC activity was enhanced (Fig. 16A&B). Based on all the data we collected, we thought working in a heme-limiting condition was unmasking the MC response, or in other words, doing experiments where heme was plentiful might have attenuated our measurement of the MC response.

We were aware that fluorescent proteins (FPs) are dependent on oxygen to fluoresce. To make sure we could still rely on the live-cell reporter in 1% oxygen conditions, I performed a time-course where I tracked the eYFP vs mCherry intensity over time for 11 cells (Fig. 17. Only data for 6 cell shown). This experiment demonstrated that even though 1% oxygen decreased the fluorescence intensity for both FPs, this effect was perfectly linear (Fig. 17). This meant that we could trust the live-cell reporter if we imaged each plate of cells after the same amount of time post induction. Another way to test this is to stop translation after low oxygen treatment and

expose the cells to 21% oxygen before imaging, allow the FPs enough oxygen to generate a fluorescent signal. The data showed that the reporter construct with no pri-miRNA hairpin (empty reporter), had increased processing in 1% oxygen. The “processing activity” was even higher without cyclohexamide (CHX). This clearly showed that the reporter was not measuring pri-miRNA processing, but rather it was likely affected by low oxygen exposure. The pri-miR-9-1 reporter with CHX did show some increased processing in 1% oxygen, on the order of what I measured with the empty reporter. I repeated this experiment and included the pri-miR-9-1-M1 reporter which contains a pri-miR-9-1 mutant that does in not process in cells. Similar to the first experiment, I saw increased processing with the pri-miR-9-1-M1 reporter. Together these results suggest our method of measuring MC processing activity in low oxygen is not good enough to confidently conclude.

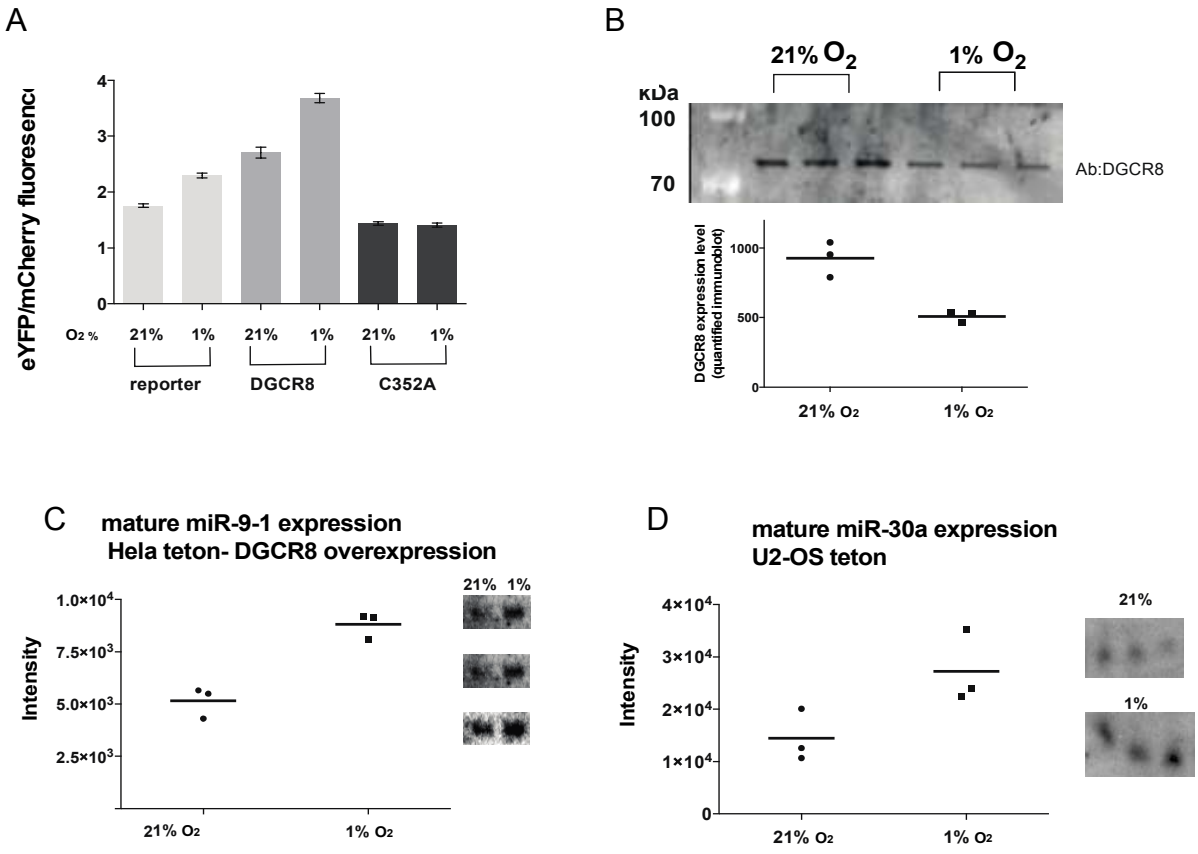
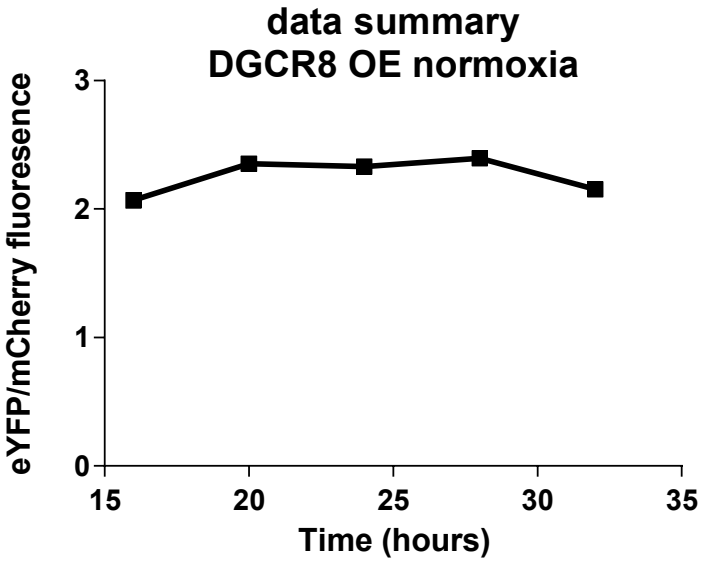


Figure 14. MC processing activity is stimulated by low oxygen. A) Live-cell assay in Hela-Tet-on cells with the pri-miR-9-1 reporter. B) Western blot of DGCR8 protein expression in 21% and 1% oxygen and quantification. C-D) Quantification of northern blot analysis with miR-9-1 probe in Hela-Tet-on cell and miR-30a probe in miR-30a probe.

A



B

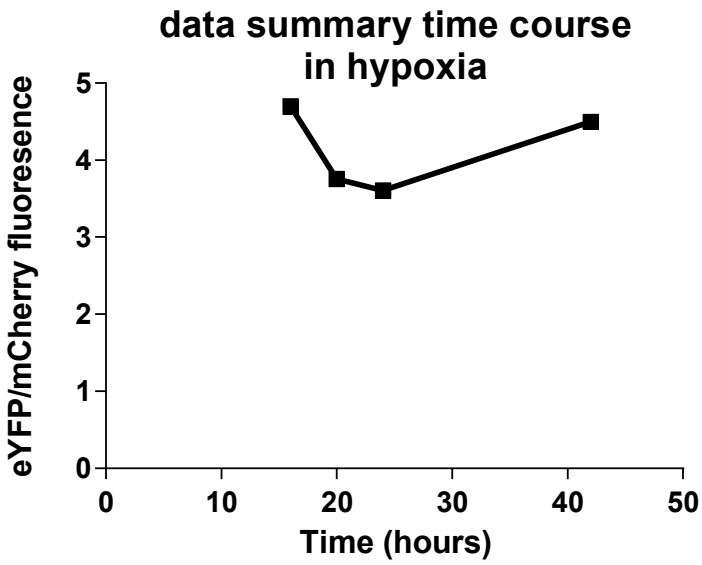


Figure 15. Time-dependence of eYFP vs. mCherry slope in normal and low oxygen conditions. A) eYFP vs. mCherry slope over time in 21% oxygen. B) eYFP vs mCherry slope in 1% oxygen. All conditions for A& B used DGCR8 over expression.

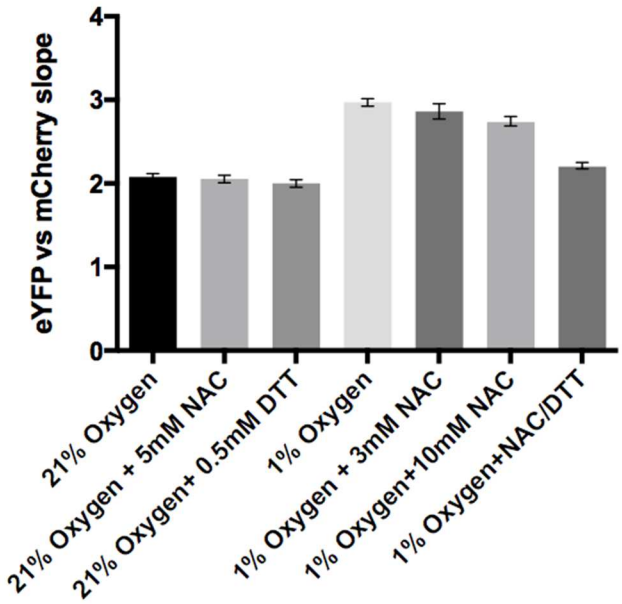


Figure 16. ROS scavengers NAC and DTT attenuate the MCs enhanced activity in low oxygen.

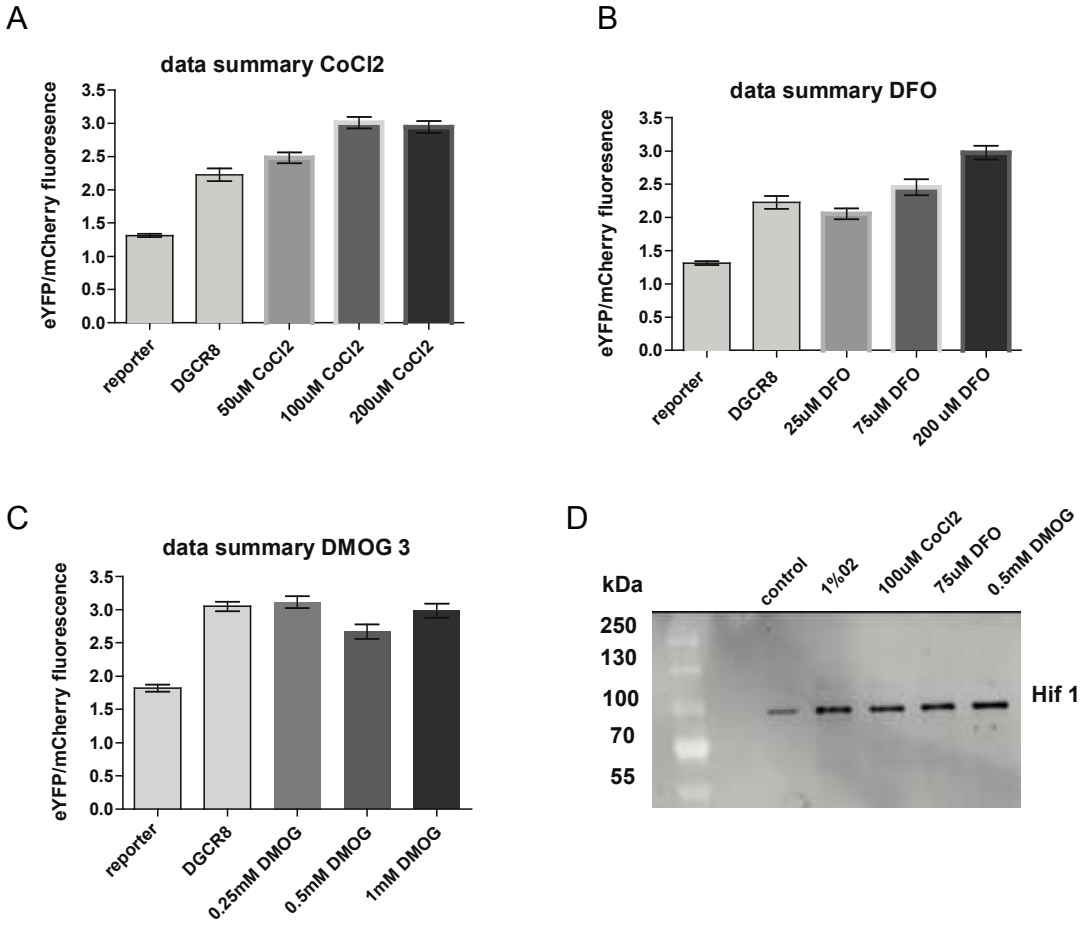


Figure 17. Chemical induction of hypoxia. A) Live-cell assay with $CoCl_2$ titration. B) Live-cell assay with DFO titration. C) Live-cell assay with DMOG titrations. D) Western blot shows induction of HiF-1α in all treatment conditions.

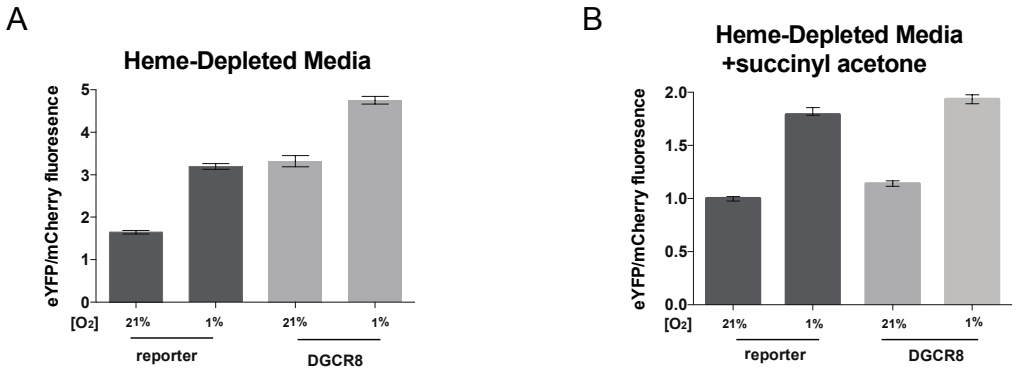
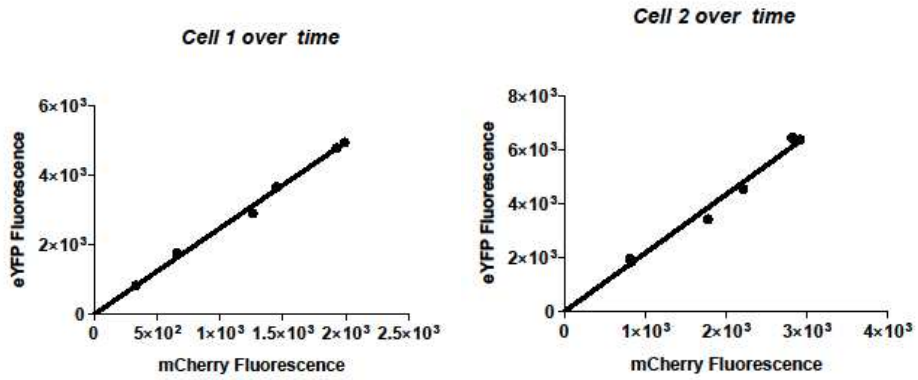


Figure 18. Heme depletion enhances MC activity in 1% oxygen. A) Live-cell assay in 21% and 1% oxygen in heme-depleted media measuring the endogenous MC and overexpression of DGCR8. B) Live-cell assay in 21% and 1% oxygen in heme-depleted media with the heme synthesis inhibitor, succinyl acetone.



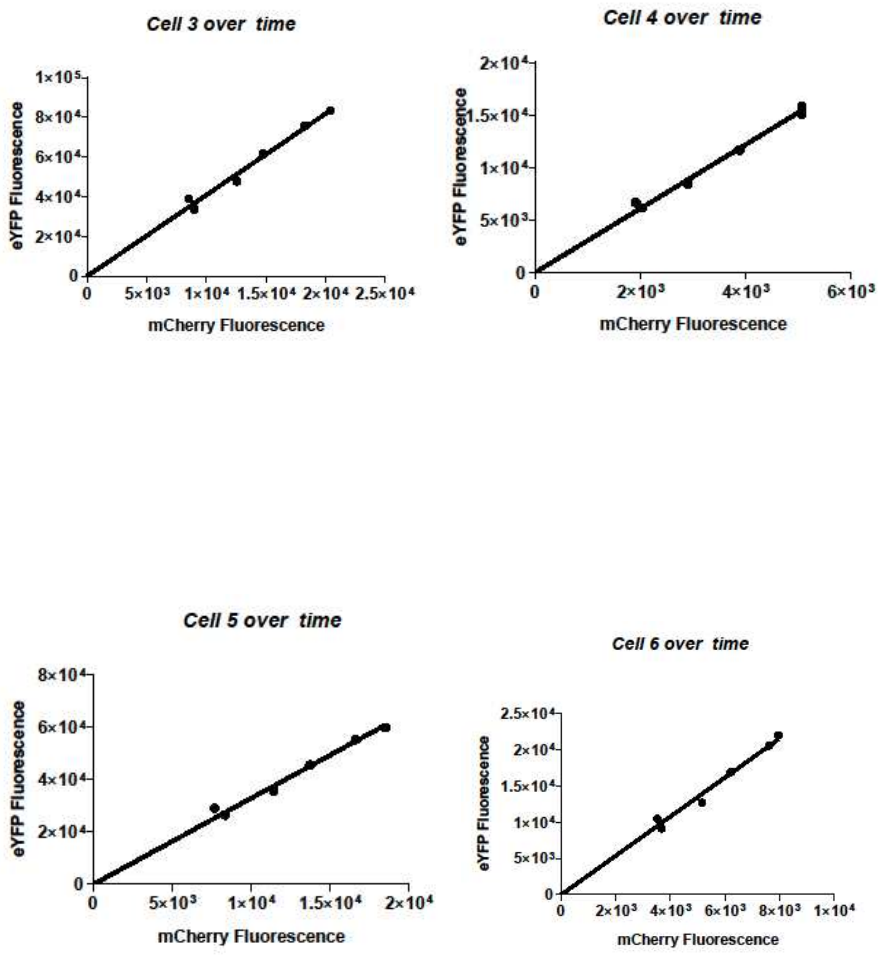
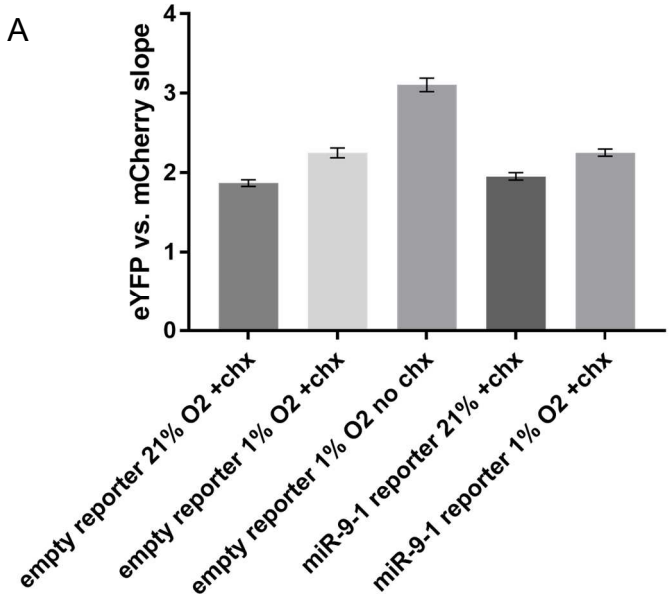


Figure 19. eYFP and mCherry signal remains linear with 1% oxygen.



B

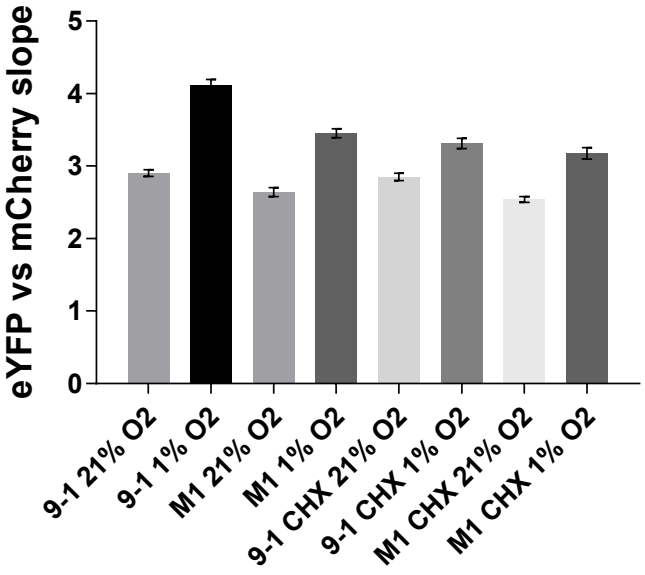


Figure 20. Live-cell assay with cyclohexamide.

A & B) Live-cell assay in 21% and 1% oxygen with and without cyclohexamide treatment shows the reporter is not measuring pri-miRNA processing.

Alkaline Hydrolysis

This procedure provides a gel electrophoresis "ladder" of hydrolyzed RNA fragments. In the procedure, 3 different hydrolysis times are used. After the experiment, select the ladder that provides the best distribution of nucleic acids over the range of lengths needed for your experiments.

Reagents Required

- 0.1–3 µg end-labeled RNA (see the 5' end labeling procedure or the 3' end labeling procedure for protocols for end-labeling RNA).
- Yeast RNA (10 mg/ml; e.g., Cat #7118 or the Yeast RNA supplied with Ambion's RNA Grade Ribonucleases)
- 1X Alkaline Hydrolysis Buffer (e.g., 50 mM Sodium Carbonate [NaHCO3/Na2Co3] pH 9.2, 1 mM EDTA; or the Alkaline Hydrolysis Buffer supplied with Ambion's RNA Grade Ribonucleases)
- Acrylamide Gel Loading Buffer (e.g., 95% Formamide, 18 mM EDTA, 0.025% SDS, 0.025% Xylene Cyanol, 0.025% Bromophenol Blue; or Gel Loading Buffer II (Cat #8546G) supplied with Ambion's RNA Grade Ribonucleases)
- 6–20% acrylamide/7 M urea sequencing gel (the percentage of acrylamide will depend on the length of RNA to be analyzed; gel size will depend on the resolution required)

Protocol

Before use, the Inactivation/Precipitation Buffer should be completed by addition of 3.2 mL of 100% ethanol.

1. Mix 0.1-3 µg of end-labeled RNA and 3 µg of yeast tRNA in a volume not to exceed 5 µl.

2. Add sufficient 1X Alkaline Hydrolysis Buffer to bring the final volume to 15 µl.

3. Aliquot 5 µl of the RNA/buffer mixture into 3 tubes labeled 1–3.

4. Heat the tubes to 95°C.

5. After 2 min, remove Tube #1 to an ice bucket.

6. After 5 min, remove Tube #2 to at an ice bucket.

7. After 15 min, remove Tube #3 to at an ice bucket.

8. Add 10 µl of Acrylamide Gel Loading Buffer to each of the 3 tubes. For an untreated control, mix 1 µl of 5' end-labeled RNA with 8 µl of Acrylamide Gel Loading Buffer.

9. Fractionate 3 µl of the RNA from all 4 tubes using a 6–20% acrylamide/7 M urea sequencing gel.

10. Use autoradiography to visualize the fractionated RNA products.

5' End-labeling RNA

***adapted from RNA lab manual by RIO, ARES, HANNON, NILSEN

Phosphatase treatment – Antarctic phosphatase

You will need to ~100pmole of 5'ends (uM=pmol/ul)– set the following reaction up in 1.5 ml tubes—everything you use should be RNase-free

And MQ water to 20 µl

2 µl of 10x reaction buffer

RNA (calculated)

2µl Antarctic phosphatase

Incubate this at 37°C for 30 minutes

Heat inactivate the enzyme at 65^{o}C for 5 minutes

Now you are ready to label with ɣ-ATP and T4 polynucleotide kinase

You will need ~20pmol of 5'ends for this reaction

4µl of MQ water

4uL of the dephosphorylated RNA stock

2µl of 10x T4 kinase buffer

8 µl of ɣ-ATP (>7000ci/mmol 167 mCI/ml) to make to compensate for the low [ATP]

2 µl T4 PNK (10 u/µl)

Total volume is 20 µl

Incubate this mixture for 1 hour at 37^{o}C

Stop the reaction with 2µl 0.5M EDTA pH 8.0

Now G25 illustra microspin columns can be used to retrieve labeled RNA or you can choose to gel purify

They recommend a load volume of 25-50 to extract labeled oligos

Resuspend the resin in the columns by vortexing

Remove the bottom

Place in collection tube

Spin at 735xg (~3000 rpm for eppindorf tubes) for 1 minute

Slowly apply sample to the apex of the mound of resin

Elute for 2 minutes at 735xg

*An additional Ethanol precipitation step is needed here as well to ensure everything is super clean

Adjust volume to 200µl (add 170 µ) with MQ water

Add 2 µl of 3M sodium acetate

Add 3 volumes of cold EtoH

Chill for 30 minutes at -80^{o}C

Centrifuge top speed for 10 minutes

Aspirate the supernatant

REPEAT – this stuff needs to clean

Resuspend in 10/µl MQ water.

***For end label of low molecular weight ladder (affymetrix10-100nt)

From manual : "For 5´-end labeling reaction, use 1 µl (50 ng each fragment) of Marker with Optikinase™ (USB PN 78334) and [γ-^{32}P]-ATP or [γ-^{33}P]-ATP."

No dephosphorylation necessary – do a 20 ul reaction with 1 ul ladder and purify with G25 column no etoh ppt needed

Phenol-Chloroform-Isoamyl Alcohol extraction (PCA)

Bring sample volume to 200ul

- If the [protein] is relatively high you should probably do a proteinase K digestion. Add proteinase K to 100ug/mL and incubate for 20 minutes at 42 degrees C → when it doubt DO THIS STEP

Add carrier to appropriate concentration

- you need a carrier if your [RNA] is < 10ug/ml (use tRNA @ 10 mg/ml, glycogen @ 20mg/ml, and glycoblue 10mg/ml).

Add 3M NaOAc pH 5.2 to final concentration of 0.3M

Add 2 Volumes of PCA and VORTEX vigorously for 1 minute

Centrifuge for 2 minutes at max speed

Remove the aqueous layer into a fresh tube

Continue on with the ethanol precipitation

Ethanol Precipitation

Add 3M NaOAc pH 5.2 to final concentration of 0.3M

Add 3 volumes of 100% filtered ethanol

Vortex to mix thoroughly

Freeze samples completely in liquid nitrogen

Thaw and centrifuge for 10 minutes at max

Remove the liquid

If there is a lot of salt you can consider doing a 70% ethanol wash…..if your RNAs are short try 85% ethanol wash → you can always do another 100% ethanol extraction

For footprinting:

Resuspend in 200uL 2mM EDTA pH 8.0 and precipitate again with 0.3M NaOAc pH 5.2 and 3 volumes ethanol

→ if its late and your ready to leave this a good storage point → store samples in freezer overnight and don't worry RNA in ethanol is just fine overnight

If you're still good to go:

Freeze in liquid nitrogen and thaw

Spin for 10 minutes at max speed

Remove liquid as completely as possible and let air dry 5-10 minutes (don't wait to long or else your RNA won't go back into solution and you will be very sad)

Resuspend in milliQ water (whatever volume you need)

SEC analysis of RNA-protein complex

Equilibrate Superdex 200 column low salt SEC buffer containing 20 mM Tris pH 8.0 and 80 mM NaCl.

Build the complex: final volume=150ul Final salt = 233mM, protein/RNA ratio 2:1

*I usually use 4uM Rhed or NC1:2uM RNA for analytical runs

1. Anneal RNA in Hepes pH 7.0 : final volume 62.7ul 65 degrees C for 3 minutes
2. Allow the RNA to cool and fold for 15 minutes at room temperature
3. While RNA is folding, mix protein with SEC (400mM NaCl, 20mM Tris pH 8.0) final volume 87.3 ul
4. Incubate at room temperature for 5 minutes.
5. Load the complex using a small disposable syringe+needle with the housing wrapped with saran wrap to guard against injecting air 100µl loop
6. Collect 0.2 mL fractions. For EM collect 0.1 ml fractions

SEC complex runs: data processing

1.Use the trace of each run to determine a volume to use as zero (usually around 5 mLs works well)

2. Create columns in Excel where you subtract each absorbance at each monitored wavelength at the volume you selected

3. Calculate [RNA]: (260nm-background-450nm-background* 260nm/450nm protein)/ E260

4. Calculate [Heme]: 450nm-background/E 450nm

5. Then create a column with the [Heme]/[RNA] ratio

*Sometimes you may be stuck using the 260/280 ratio of the protein, RNA and protein-RNA complex to estimate the protein-RNA ratio → this isn't very accurate

((260/280 RNA-260/280 complex)*E260) / (260/280 RNA * (260/280 complex-260/280 protein)*E280)

6. Put the background subtracted wavelength columns in graphpad to display your curve (use 450 nm and 260 nm – no need to display 280)

7. Put the column of whatever ratio you want to display in the spreadsheep and snap it to a right Y- axis. Delete regions that are not overlaying your SEC curves unless they are informative.

Preparing Uranyl Acetate Negative Stain (from the Rome Lab - UCLA)

Standard = 4% w/v stock:

4g Uranyl Acetate Powder

Heat water in small flask until boiling

Re-hydrate with enough boiling water to get powder into solution

Let this solution cool then bring volume up to 100mLs

Dilute your solution to 1% with water and filter through 0.22micron filter before use.

Staining your sample

Cool a metal heat block on ice. Right before you begin staining, cover with parafilm

Glow discharge grids

Spot 10ul of each sample for the grids on the parafilm on the chilled heating block

Using EM grade negative-action tweezers to float your grid on the sample (carbon-coated side down)

Incubate the grid in the sample for 5 minutes

While you wait, Spot 20-100 ul of stain on the chilled heatblock

Gently remove the grid from the sample and blot from the edge with filter paper

Place grid on the stain for 5 minutes to stain

Wick away excess stain from the side again and store face-up in a glass-covered dish to dry

EM imaging procedure for T-12 or TF-20

>Go to Lo mag (LM500) & removed the objective aperature

*the objective aperture is the center aperature – you move the aperatures in and out with the small metal lever at the bottom. Rotate right takes it out of the beam

> expand the beam to fit the whole screen and center it with the trackball (left controller)

> focus the beam with the intensity knob on the larger of the circles on the screen

> increase magnification to ~ 970(condense beam then increase) and put the objective aperature back in - center the objective manually (small ring is axial movement and side ring is perpendicular movement). Expand the beam using the intensity knob and make sure it expands concentrically. Concentric expansion means the objective is aligned.

>scan the grid at low magnification looking for an area you would like to image

>go to your working magnification ~52000x for our small complex, find some dirt near an area you want to image and focus on it

>Focusing: press eucentric focus – then put in the small viewing screen and the ocular lenses and fine-tune your focus with the focus knob - you should be seeing particles. You may have to adjust the intensity so you can see.

>Increase the magnification a little and put the area of interest at eucentric height manually using the wobbler and the z axis +/- knobs (OR you can change the angle to -30 then use +/- to bring your object back into center)

> go back to your working magnification – focus again- (press eucentric focus and manually fine tune the focus using the small viewing screen and ocular lenses

> take a look at your work (press “view” in digital micrograph). Collect a live FFT and make sure the astigmatism is correct (perfectly circular thon rings). Fine tune focus if you need to and take a image “start aquire”.

>Save image!!!

RNAse Protection Assay

Phosphate Footprint buffer (PFB) = 600mM NaCl, 20mM Na2HPO4/NaH2PO4 pH 7.0

Final NaCl = ~100mM

[RNA] needs to be low to ensure 100% binding

Preparation for 4 conditions

Anneal :

RNA 25kCPM

In 10mM Hepes pH 7.0

----------79.2ul 65 degrees C for 3 minutes

Bind protein – concentration determined by measure K_d

X protein

PFB buffer

---------- 15.8ul let bind 10 minutes at room temperature

Disperse into 20ul fractions for each respective condition:

Sample list:

Add 10ug of carrier RNA to each tube

Add enzyme and mix well-you can play with incubation time/temp/ and enzyme concentration. Enzyme works in temperatures from 55-4 degrees Celcius

add 0.01, 0.02 units of T1 and 0.04, 0.08 units of V1

incubate desired temperature for 15 minutes:

Conditions used:

Add 3ul of 5mg/ml ATA and mix well – *if the protein concentration is low (below 400nM) – it's a great move to add an equal amount of 2x RNA loading buffer, normalize the counts you are going to load with the scintillation counter and run your gel-→ no need to extract!*

Dilute reactions with 200ul of SDS extraction buffer, add 0.3M NaOAc and extract with two volumes of phenol chloroform and then 3 volumes of 100% ethanol

Re-suspend in 10ul water and add 10 ul 2x RNA loading buffer

Load 8 ul onto gel to visualize

Dry gel and image with phophor screen

References

1. Celander DW. Probing RNA structures with hydroxyl radicals. Curr Protoc Nucleic Acid Chem. 2001;Chapter 6(Ii):Unit 6.5.

2. Wang L, Chance MR. Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. Anal Chem [Internet]. 2011;83(19):7234–41.

3. White R, Crow J, Spear N, Thomas S, Patel R, Green I, et al. Making and working with peroxynitrite. Methods Mol Biol. 1998;100(5):215–30.

4. Faller M, Toso D, Matsunaga M, Atanasov I, Senturia R, Chen Y, et al. DGCR8 recognizes primary transcripts of microRNAs through highly cooperative binding and formation of higher-order structures. RNA [Internet]. 2010 Aug [cited 2013 Nov 14];16(8):1570–83.

5. Weitz SH, Gong M, Barr I, Weiss S, Guo F. Processing of microRNA primary transcripts requires heme in mammalian cells. Proc Natl Acad Sci U S A. 2014;111(5):1861–6.

6. Bandara K V, Michael MZ, Gleadle JM. Hypoxia represses microRNA biogenesis proteins in breast cancer cells. BMC Cancer [Internet]. 2014;14(1):533.

7. Wu C, So J, Davis-Dusenbery BN, Qi HH, Bloch DB, Shi Y, et al. Hypoxia Potentiates MicroRNA-Mediated Gene Silencing through Posttranslational Modification of Argonaute2. Mol Cell Biol. 2011;31(23):4760–74.

8. Clanton TL. Hypoxia-induced reactive oxygen species formation in skeletal muscle. J Appl Physiol. 2007;102(6):2379–88.

9. Clanton T. Yet another oxygen paradox. J Appl Physiol. 2005;99(4):1245–6.

10. Du J, Xu R, Hu Z, Tian Y, Zhu Y, Gu L, et al. PI3K and ERK-induced Rac1 activation mediates hypoxia-induced HIF-1?? expression in MCF-7 breast cancer cells. PLoS One. 2011;6(9):1–9.

11. Kováčik J, Klejdus B, Bačkor M. Nitric oxide signals ROS scavenger-mediated enhancement of PAL activity in nitrogen-deficient Matricaria chamomilla roots: side effects of scavengers. Free Radic Biol Med [Internet]. 2009;46(12):1686–93.

12. Yuan Y, Hilliard G, Ferguson T, Millhorn DE. Cobalt inhibits the interaction between hypoxia-inducible factor-alpha and von Hippel-Lindau protein by direct binding to hypoxia-inducible factor-alpha. J Biol Chem [Internet]. American Society for Biochemistry and Molecular Biology; 2003 May 2;278(18):15911–6.

13. Woo KJ, Lee T-J, Park J-W, Kwon TK. Desferrioxamine, an iron chelator, enhances HIF-1alpha accumulation via cyclooxygenase-2 signaling pathway. Biochem Biophys Res Commun [Internet]. 2006 Apr 28 [cited 2015 Sep 23];343(1):8–14.

14. Mace KA, Yu DH, Paydar KZ, Boudreau N, Young DM. Sustained expression of *Hif-1α* in the diabetic environment promotes angiogenesis and cutaneous wound repair. Wound Repair Regen [Internet]. Blackwell Publishing Inc; 2007 Sep 1 4];15(5):636–45.

15. Barr I, Weitz SH, Atkin T, Hsu P, Karayiorgou M, Gogos JA, et al. Cobalt(III) Protoporphyrin Activates the DGCR8 Protein and Can Compensate microRNA Processing Deficiency. Chem Biol [Internet]. Elsevier; 2015;22(6):793–802.