REGULATORY MODIFICATIONS OF
MICRORNA AND SHORT HAIRPIN RNA PRECURSORS

Robert Lanning Sons

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cell and Developmental Biology, in the Department of Cell Biology and Physiology in the School of Medicine.

Chapel Hill
2015

Approved by:

Scott Hammond

Pat Brennwald

John Reader

Bill Marzluff

Zbig Dominski

ABSTRACT

Robert Lanning Sons: Regulatory Modifications of MicroRNA
and Short Hairpin RNA Precursors
(Under the direction of Scott Hammond)

RNA-interference (RNAi) is ubiquitously used as a research tool, with dozens of candidates using this technology currently in FDA clinical trials. A particular segment of the RNAi world, short hairpin RNAs (shRNAs) seeks to enable stable, long-term reduction of a target protein. Here, we present an optimization of shRNA structure that allows increased reduction of targeted proteins as unwanted processing events are bypassed, increasing the amount of mature RNAi species created. In Chapter 1, we present current literature as it pertains to RNAi biogenesis, regulation, and optimization of shRNA effectiveness. In Chapter 2, we present our findings of extensive shRNA precursor degradation, and complete an unbiased screen for structures that bypass this degradation and optimize RNAi activity. In Chapter 3, we offer our observations of negative processing events associated with RNA Polymerase III (RNAPIII)-driven shRNAs, including but not limited to the same extensive precursor degradation. Together, these data present new details in the dysregulation of endogenous miRNA biogenesis by shRNAs, and offer guidance towards more accurate and potent engineered shRNA regulators less likely to induce off-target effects.

Dad, I wish you were still here.

ACKNOWLEDGMENTS

PREFACE

Chapter 1, a literature review, entitled "RNAi Biogenesis and Application", was written for this thesis, as was Chapter 3, a research article entitled "Defects in First Generation shRNA Design, and Impact on Endogenous Precursor miRNAs". Chapter 2's figures and text are incorporated from an article under review in PLOS ONE entitled "Optimizing Short Hairpin RNA Design to Maximize RISC Loading". My dissertation work has also included a collaboration with the Hornstein lab at the Weizmann Institute of Science, which enabled normalization of precursor-specific deep sequencing, revealing inhibition of Dicer activity in amyotrophic lateral sclerosis. In addition to this authorship, which was recently accepted to The EMBO Journal, I am a co-author on two other collaborations with the Linnstaedt and Deshmukh labs that are currently in review at PAIN and Oncotarget. Together, these projects and publications represent my dissertation efforts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ac-pre-miRNA   Ago2-cleaved precursor microRNA

Ago   Argonaute

AgoshRNA   alternatively processed shRNA

dsRBP   double-stranded RNA binding protein

dsRNA   double-stranded RNA

miRNA   microRNA

mRNA   messenger RNA

PAPD4/PAPD5   poly(A) polymerase PAP-associated domain containing 4 or 5

PARN   poly(A)-specific ribonuclease

pre-miRNA   precursor microRNA

pri-miRNA   primary microRNA

RISC   RNAi-induced silencing complex

RNAi   RNA interference

RNAPII   RNA Polymerase II

RNAPIII   RNA Polymerase III

shRNA   short hairpin RNA

shRNAmir   miRNA-like shRNA

siRNA   short interfering RNA

sshrna   short shRNA

TUTase   terminal-uridyl transferase

XPO-5   Exportin-5

Chapter 1: RNAi Biogenesis and Application

Overview

This chapter's purpose is to introduce the reader to the mechanism and regulation of

RNAi, and then show how researchers use that knowledge to study and perturb gene

function. As an introduction, the broad strokes of RNA interference (RNAi) what it does, and

how it does it, will be presented. The second main section, biogenesis, describes the pathway

and its regulators, beginning with transcription of what will become a mature RNAi

molecule, ending with what is known about how a target mRNA is repressed, and including

descriptions of the changing RNA secondary structure throughout. Lastly, researchers'

harnessing of RNAi will be discussed, including the strengths and weaknesses associated

with the main variations in input structure. Lessons regarding biogenesis preferences and

regulation are relevant for scientists across the range of basic to clinical research.

RNAi Activity

RNA interference (RNAi) was first discovered in *Caenorhabditis elegans* through

humble beginnings; ingestion or injection of double-stranded RNA (dsRNA) resulted in non-

stoichiometric knockdown of complementary messenger RNA (mRNA), pointing to the

possibility of catalytic activity (Timmons and Fire 1998; Fire et al. 1998). Endogenous use of

this regulatory pathway occurs via 21-23nt single-stranded RNA called microRNAs (miRNAs) that are conserved and clustered across most eukaryotic species. These miRNAs transit a biogenesis pathway from an initial transcript, the primary miRNA (pri-miRNA) hundreds of nucleotides (nts) long, to precursors (pre-miRNA) ~70nts long, to the RNA:RNA duplex containing a mature molecule that ends up in complex with an RNAi-induced silencing complex (RISC), through which cells can regulate protein levels (Altuvia 2005; Bartel 2004, 2009). Within the RISC complex, the miRNA is bound by the core RISC protein, Argonaute (Ago). Once bound by Ago, regulation of targeted mRNAs depends on complementarity between the bound miRNA and the target mRNA. Though the strongest activity associates with the highest complementarity between guide:target, imperfect complementarity can still induce RNAi, such that a single miRNA could promiscuously regulate ~100 target sites (Zamore et al. 2000; Jackson et al. 2003; Brennecke et al. 2005; Farazi et al. 2008; Doench and Sharp 2004; Gu et al. 2014).

mRNA cleavage, degradation after decapping and deadenylation, and translation repression are all observed consequences of RNAi induction, but there was confusion as to which was the main contributor to reductions in protein levels. However, the portion of the guide strand that most contributed to RNAi was honed in upon: complementarity between nucleotides 2-9, counting from the 5' phosphate of the guide strand, and a target mRNA, was the minimal amount of base-pairing necessary to induce RNAi. This heptamer base-pairing requirement enables the promiscuity that places >30% of protein coding genes within regulatory control of miRNAs (Lim et al. 2005; Lewis et al. 2005; Bartel 2004; Aravin and Tuschl 2005; Bartel 2009; Behm-Ansmant et al. 2006; Pillai et al. 2007).

More recently, mRNA degradation was revealed to be the dominant mechanism of RNAi activity. Most protein levels resist large miRNA-mediated reductions in their quantity; hundreds of proteins "fine-tuned" in their expression levels are more to be expected. However, more than 60% of human protein-coding genes have maintained miRNA target sites above background levels, acknowledging the value of this fine-tuning (Baek et al. 2008; Selbach et al. 2008; Friedman et al. 2009; Guo et al. 2010). Cleaved mRNA is subsequently oligo-uridylated (1-35nt uridine 3' tail), accompanied by 5' decapping and 3' degradation (Shen and Goodman 2004). Translational repression of mRNAs via RNAi occurs via mechanisms still being elucidated, but one option includes mRNA sequestration to P-bodies, in addition to potential decay of mRNA at that location (Liu et al. 2005; Gu and Kay 2010; Huntzinger and Izaurralde 2011).

Post-transcriptional regulation doesn't only affect mRNA, it affects mature miRNAs as well; their quantities can be disproportionate to pri-miRNA levels, proteins can bind pre-miRNAs, accelerating production of the mature strand, and ubiquitously expressed pri-miRNA transcripts can be blocked from producing their mature miRNAs except in a tissue-specific fraction of backgrounds (Eis et al. 2005; Trabucchi et al. 2009; Obernosterer 2006; Thomson 2006). Even increased cell-cell contact can affects microRNA biogenesis by enhancing Drosha activity, as well as overall RNAi activity through more efficient formation of RISC (Hwang et al. 2009).

A major driving force within the field of miRNAs has been the discovery of their dysregulation in multiple diseases, most prominently cancer. Impaired biogenesis enhances cell transformation and tumorigenesis, and on the diagnostic/prognostic side, miRNA quantitation can discriminate between more and less aggressive subtypes of cancer. Multiple

factors result in miRNA dysregulation, including reductions in levels of biogenesis machinery proteins. Overall, global miRNA expression is lower in cancer vs. normal tissues, such that globally higher levels of miRNA expression equate to a more differentiated cell state (He et al. 2005; Lu et al. 2005; Blenkiron et al. 2007; Kumar et al. 2007; Chang et al. 2008).

The use of the RNAi pathway has been incredibly influential within the scientific community, with over 70,000 papers referencing siRNA or shRNA in a rudimentary search of PubMed, dozens of clinical trials currently in Phase I/II, and a first for the field, completion of Phase III for an RNAi therapeutic expected in 2016 (Haussecker 2015; Kubowicz et al. 2013; Haussecker and Kay 2015; PRNewswire 2015). However, the dangers of dysregulated biogenesis of miRNAs as summarized previously are stark reminders of the delicacy that must be employed if the RNAi pathway is to be effectively deployed on researchers' and clinicians' behalves. Below, the components of biogenesis will be presented, prior to concluding this chapter with a discussion of artificial RNAi tools and their intended and unintended interactions with biogenesis machinery.

Biogenesis

The immediately subsequent sections will describe the journey of an endogenous pri-miRNA transcript through the microprocessor complex, consisting of Drosha and DGCR8, as its' 5' and 3' single-stranded RNA flanks are cleaved to leave behind a ~70nt pre-miRNA, the canonical stem-loop hairpin. The precursor is then transported to the cytoplasm by Exportin-5 (XPO-5), for subsequent cleavage of the loop by Dicer, resulting in a ~21nt double-stranded RNA (dsRNA) duplex, one of which is the mature miRNA, or guide strand,

4

the other which is a passenger strand, intended to be degraded instead of incorporated into RISC (Fig. 1.1). This primer on biogenesis will conclude after a discussion of post-transcriptional modifications and their effects on biogenesis regulation, followed by a presentation of how researchers have harnessed RNAi.

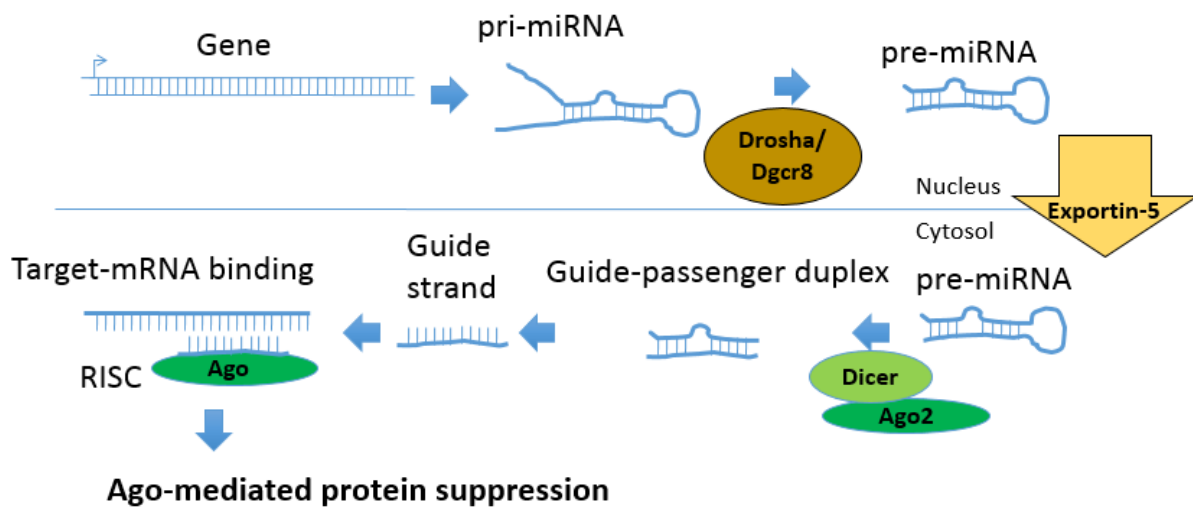**Figure 1.1: Canonical miRNA Biogenesis**. Drosha and DGCR8 measure the secondary structure, cleaving the 5'/3' flanks to produce a pre-miRNA with thermodynamically unstable stem ~19 nucleotides in length. The pre-miRNA transits the nuclear pore complex via Exportin-5, after which Dicer cleaves the loop, producing a dsRNA guide-passenger duplex. The guide strand is bound by Ago, allowing RNAi activity.

Microprocessor Complex: Drosha/DGCR8 production of pre-miRNA

Pri-miRNA transcripts originate in the nucleus, and can vary in length, interaction with splicing machinery, polyadenylation, and be independent transcripts or lie within other mRNAs. The presence of a precursor structure within the 3' UTR of another mRNA or an intronic region only moderately inhibits the protein production from that parental mRNA (Bracht 2004; Cai 2004; Kim and Kim 2007; Morlando et al. 2008; Melamed et al. 2013). Drosha and DGCR8 form the Microprocessor complex that cleaves the 5' and 3' flanking sequences of the pri-miRNA to produce a pre-miRNA hairpin ~70nt in length with a 5' phosphate and a 2nt 3'-OH overhang. (Denli et al. 2004; Landthaler et al. 2004; Gregory et al. 2004; Han 2004; Nguyen et al. 2015; Basyuk et al. 2003; Lee et al. 2003). Two DGCR8 proteins measure 22 basepairs from the hairpin loop; ~97% accurate and precise Drosha cleavage occurs if the 5' and 3' ssRNA flanks are 11nt away from where these DGCR8 protein measure (Han et al. 2006; Ma et al. 2013). Both secondary structure and primary sequence can work in concert with multiple proteins, including hnRNP A1, SMAD, p53, and Drosha, to enhance cleavage and production of pre-miRNA (Guil and Cáceres 2007; Michlewski et al. 2008; Davis et al. 2008, 2010; Suzuki et al. 2009; Auyeung et al. 2013). Of particular interest once harnessing of RNAi, and how miRNA-like features enhance our ability to induce RNAi, will be discussed, pri-miRNAs that are not cleaved by Drosha concurrent with their transcription are less efficiently processed to mature miRNA form (Pawlicki and Steitz 2008).

Exportin-5: the regulator of nuclear pre-miRNA export

Exportin-5 (XPO-5) is a karyopherin that includes among its cargo pre-miRNA, for export through the nuclear pore complex into the cytoplasm, in a Ran-GTP dependent, saturable manner. XPO-5 prefers >16nt of dsRNA in the stem, and the 5' phosphate and 3'-OH overhang that result from Drosha processing. However, it can accommodate a 3' blunt end, has no loop preference, but a 5' overhang is inhibitory (Yi 2003; Bohnsack et al. 2004; Lund 2004). In addition to transport, XPO-5 may act as a reservoir for pre-miRNA, protecting them from nuclease cleavage (Zeng and Cullen 2004). In addition to competition amongst precursors, another cargo for XPO-5 is Dicer mRNA; overexpression of pre-miRNAs results in reduced Dicer levels (Bennasser et al. 2011). Inactivating mutations in cancer traps precursors in the nucleus, reducing biogenesis and RNAi activity; additional expression of XPO-5 reverses precursor segregation, biogenesis and activity, and produces tumor-suppressing effects (Sonia A Melo 2010).

Dicer

Prior to becoming a single stranded guide miRNA, the loop and passenger strand must be removed. The canonical pathway consists of cytoplasmic Dicer that cleaves the dsRNA from the loop using two RNAseIII domains, producing a dsRNA that now has two 3'OH overhangs. The Paz domain, also present in Ago2, binds both free ends of the precursor. (Hutvágner et al. 2001; Knight 2001; Song et al. 2003; Lingel et al. 2004; Zhang et al. 2004; Gurtan et al. 2012). A dsRNA binding domain contains a non-canonical nuclear localization signal, and though Dicer has been reported in the nucleus, evidence of its

contribution to miRNA biogenesis and RISC maturation have been limited to the cytoplasm (Doyle et al. 2013; White et al. 2014; Gagnon et al. 2014). Lastly, Dicer's helicase domain assists in transitioning thermodynamically unstable substrates to RISC (Soifer et al. 2008). Two double-strand RNA-binding proteins (dsRBPs), TRBP and PACT, interact with Dicer and are required for precursor maturation into RISC. They balance each other; TRBP inhibits protein kinase R (PKR) and its downstream innate immune response, while PACT activates PKR (Chendrimada et al. 2005; Kok et al. 2007).

Once the precursor is cleaved, the RNA duplexes unwind, with the mature miRNA/guide strand entering into complex with Ago2, and the passenger strand being destroyed. This loading happens in an asymmetric manner based upon thermodynamics: the 5' end with weaker binding relative to the opposite strand of the RNA duplex is the strand destined to become the guide. Dicer makes this comparison of the ~4 basepairs of each end; the RNA duplexes shift within for sensing of thermodynamic asymmetry (Schwarz et al. 2003; Chendrimada et al. 2005; Kok et al. 2007; Noland et al. 2011).

In addition to sensing thermodynamic stability, Dicer recognizes changes in miRNA precursor structure, altering its efficiency and positioning of cleavage. Both the 5' and 3' ends of precursor miRNAs are recognized by Dicer; +/-1nt differences in length of the 3' overhang are slightly tolerated; +/-2nt changes in the overhang length are sub-optimally processed. Additionally, Dicer prefers a precursor with a base-paired 5' end and a large loop, but outside of these restrictions, is flexible in terms of potential substrates (Park et al. 2011; Feng et al. 2012). Within the stem, Dicer ignores asymmetric bulges; i.e. a single nt extra bulge on the 3p strand results in a 3p mature miRNA 1nt longer than if the bulge were

symmetric, contributing to the length and positioning diversity observed with some endogenous microRNA guide strands (Starega-Roslan et al. 2011).

RNAi-induced Silencing Complex (RISC)

After thermodynamic sensing and comparison of the base-pairing strength of 5' ends of the guide:passenger duplex by Dicer, the guide is chosen and incorporated into the catalytic engine of RNAi, RISC. RISC was discovered in *Drosophila* to contain an RNA-directed nuclease and a guide RNA, later identified as Argonaute 2 (Hammond et al. 2000, 2001). This guide strand can then direct RISC for cleavage of a series of complementing mRNA targets (Hutvágner and Zamore 2002; Martinez et al. 2002). The PIWI domain, similar to ribonuclease H, confers "slicing" ability on Argonaute 2 opposite the guide strand, and the PAZ and PIWI domains together form a groove for guide/target mRNA substrate binding. In mammalian cells, only Ago2, out of Ago 1-4, has this cleavage ability, and miRNAs indiscriminately associate with all four Agos (Lee et al. 2004)(Lee et al. 2004)(Liu 2004; Meister et al. 2004; Song et al. 2004). In addition to cleaving passenger strands and target mRNAs, Ago2 can also cleave the 3P of a few pre-miRNAs in the same position, forming a hairpin with a 3P ~11-12nt shorter, termed Ago2-cleaved precursor miRNA (ac-pre-miRNA) (Diederichs and Haber 2007). Surprisingly, there are a few reports of pre-miRNA guiding active RISC via Ago2 in the absence of Dicer (Tan et al. 2011, 2009). It is interesting to speculate whether an ac-pre-miRNA is the actual guiding component, not an uncleaved pre-miRNA, especially within the context of our observations described in Chapters 2 and 3.

In humans, miRNA precursors, together with Ago2, Dicer, and TRBP, are sufficient to assemble, in an ATP-independent manner, RISC capable of cleaving complementary RNA targets. Importantly, Ago2 and Dicer can be in complex, termed the miRNA loading complex (miRNP), prior to pre-miRNA substrate presence, likely coupling Dicer processing of precursor to unwinding and Ago2 loading. Once bound, the miRNP complex cannot be exchanged, and expression of mRNA complementary to an Ago-bound guide strand is positive feedback that increases the quantity of that Ago-bound guide strand, likely through increased stability (Krol et al. 2010; Maniataki 2005; Flores et al. 2014). This preservative observation makes sense, as increasing Ago levels increases mature miRNA levels, extending their half-lives, acting as a reservoir (Winter and Diederichs 2011). Though miRNA:mRNA duplexes have been found to exist in cells without Ago association, and there are 13-fold more miRNA than Ago proteins in a cell, Ago-bound miRNA is a better indicator of a miRNA's potency than overall expression (Janas et al. 2012; Flores et al. 2014).

In *Drosophila*, dsRNA with perfect stems, termed siRNAs, segregate into different Argonaute proteins. A bulge in the Ago2 cleavage site (opposite nts 9/10 from the 5' phosphate), shunts dsRNA to RISC incorporation around Ago1. This bulge/no-bulge determinant trumps the 5' thermostability asymmetry rule of sorting controlled by Dicer. Similar mismatches between a guide strand and a target mRNA similarly inhibit Ago2 cleavage of the target mRNA. Mismatches 5' or 3' of this central stem region, nucleotides 9-11, as well as G:U non-canonical basepairs also shunt a dsRNA away from Ago2. These less thermodynamically stable structures are the majority of precursor structures in mammalian miRNAs: ~80% of endogenous mouse guide:passenger duplexes contain central/seed,

central/seed/3', or seed/3' mismatches (Okamura et al. 2004, 2009; Czech et al. 2009; Kawamata et al. 2009; Förstemann et al. 2007).

This Ago preference in *Drosophila* dovetails well with what knowledge surrounding mammalian Agos; they overlap more in terms of guide strand binding, regardless of RNA duplex thermodynamic stability/stem-central bulges. Ago1 and Ago2 can both utilize perfectly matched stemmed siRNAs (even though only Ago2 can cleave and unwind them), and all Agos can utilize bulged miRNA duplexes for translational repression. Magnitude of potential repressive ability, regardless of the initial bulged or non-bulged stem nature of the input RNA:RNA duplex, remains equivalent across the four Agos (Su et al. 2009).

Post-Transcriptional Modifications and their Regulatory Impact on miRNAs

The post-transcriptional modifications and regulation most studied surround the conserved let-7 family. Lin28 binds a "GGAG" motif in the loop of pre-let-7, recruiting the terminal-uridyl transferases (TUTase) TUT-2,4,7 to the precursor for subsequent oligo-uridyl tailing (~15 nt in length). The length of the oligo-uridine tail determines what occurs next; a mono-uridyl tail results in the optimal 2-nt overhang for subsequent Dicer processing of let-7 precursors, while a longer tail becomes a substrate for DIS3L2 3'-5' exonuclease activity and inhibition of Dicer processing. This regulatory mechanism can result in a ubiquitously expressed precursor, for which its cognate mature strand is absent (Heo et al. 2008; Newman et al. 2008; Hagan et al. 2009; Heo et al. 2009; Thornton et al. 2012; Ustianenko et al. 2013; Heo et al. 2012).

Precursor miRNAs are commonly modified (5-15% of their species' deep-sequencing reads), typically via 3' monoadenylation or mono-/oligo-uridylation. Uridine tails average 5-7nt in length, and are usually associated with trimming events or concentrated more heavily with precursors containing a Lin28 binding motif. Overall, positioning and modification of the 5' is stable, while the 3' terminus is relatively unstable (Burroughs et al. 2011; Newman et al. 2011). Uridyltransferases TUT-4/7 are localized to the cytoplasm, while components of the exosome, RRP6 and DIS3, are localized to both the cytoplasm and nucleus. Ago-bound precursors have been found to be surveilled for appropriate 3'-OH overhang length by RRP6 and DIS3: RRP6 prefers <1nt overhangs, while DIS3 prefers >2nt overhangs. Together, TUT-4/7 are used to signal precursors for DIS3 degradation, fine-tuning mature expression levels (Liu et al. 2014)

Tailing, such as 3' adenylation and uridylation, is a mechanism or language of control. In *Chlamydomonas,* uridylation of mature miRNAs can induce instability and inhibit Dicer processing(Ibrahim et al. 2010). In *Arabidopsis,* the addition of 1-5 uridines to the 3' of guide strands results in degradation, particularly when a target RNA with extensive complementarity is present (Li et al. 2005; Yang et al. 2006; Ameres et al. 2010). This language of control extends to target mRNAs as well; miRNAs induce oligouridylation of complementary sequences (Lim et al. 2014).

In mammals, uridylation of guide strands can modify a target profile, releasing a target from repression, and do not predict Ago-association, removal, or degradation (Jones et al. 2009; Flores et al. 2014). Monoadenylation by the non-canonical poly(A) polymerase PAP-associated domain containing 4 (PAPD4) can reduce RISC targeting or stabilize selective species of mature miRNAs (Burroughs et al. 2010; Ambrogio et al. 2012; Katoh et

al. 2009). However, adenylation by PAPD5 can recruit the poly(A)-specific ribonuclease (PARN), resulting in mature miRNA degradation (Boele et al. 2014). This compilation of non-templated additions associated with changes in stability may not be the whole story though; some miRNA species are more stable than others without any observed changes in terms of non-templated modifications (Gantier et al. 2011).

Biogenesis pathway: enhancing vs. inhibiting RNAi usage

The previous sections describe the major known factors involved in miRNA biogenesis. We understand the main ways in which miRNAs exert influence on complementary mRNAs. The second half of this review will cover the successes and failures as we have attempted to harness RNAi. As you might imagine from the ease with which dysregulation of the biogenesis pathway results in transformation and tumorigenesis, many portions of the biogenesis pathway have been shown to be negatively affected by our attempts to artificially direct RNAi activity. The promiscuity of targeting is also an issue; when a single nucleotide change in the length/position of a guide strand can release a target from repression, or introduce a new, unintended target, understanding and optimization of biogenesis is required for trust of these tools.

We have read about oligouridylation being a language of clearance. A couple of sections below, we will discuss a common first generation tool with which the scientific community commonly harnesses RNAi, a tool currently in human clinical trials: a precursor with a 5-uridine tail, transcribed at a high level of expression. Chapter 3 will present some

negative effects we have observed, both with regard to the tool, as well as to endogenous precursors.

The more we learn about biogenesis, the more we understand why some attempts to harness RNAi did not show promise; asymmetric bulges that were previously tested were not known at the time to shift Dicer cleavage, introducing unwanted flanking nucleotides into guide strands, potentially releasing targets from optimum levels of repression. The sensitivity of the Drosha/DGCR8 ruler similarly would introduce cleavage products beneath the 97% accurate and precise product that now understood, can be obtained.

We have identified many contributors to successful RISC formation. Some RNAi harnessing seeks to skip steps of biogenesis. However, we also see transcription that when unlinked from (Drosha) processing, reduces subsequent processing efficiency. What reductions in processing efficiency are bottlenecks leading to endogenous miRNA dysregulation and toxicity? The field has moved towards how to apply RNAi to a gene of choice, with minimal disruption to the biogenesis pathway's other responsibilities. The remainder of Chapter 1 will present the best piggy-back methods currently known, and Chapters 2 and 3 will present some previously unknown factors that negatively affect tools currently in use both at the bench and in the clinic.

Harnessing RNAi

Research at the bench and the clinic has focused on the most effective and least toxic avenue for deployment of RNAi-based tools (Rao et al. 2009). Current options include short-

interfering RNAs (siRNAs), short-hairpin RNAs (shRNAs), miRNA-like shRNAs (shRNAmirs), and non-canonically processed shRNAs called AgoshRNAs. siRNAs are perfect RNA duplexes that skip directly to RISC incorporation, and are time- and cell-limited in their use due to their dependence on transfection. Plasmid and lentiviral production of the other forms of shRNA allows a large extension of potential targets, as well as the use of the wide range of regulatable RNAPII promoters. shRNAs are perfectly-stemmed shRNAs, usually transcribed by RNA polymerase III (RNAPIII) to produce solely the pre-shRNA, intended to skip Drosha processing. shRNAmirs, however, maintain a loop, and 5'/3' flanks from endogenous miRNAs, transiting the entirety of the canonical miRNA biogenesis pathway. Some maintain bulges in their stems, but this structural modification does not describe the majority in use. AgoshRNAs have shorter stems than normal, a smaller loop, and transit a Drosha- and Dicer-independent pathway on their way to RISC formation described at the conclusion of this chapter.

Expanding beyond single-target knockdown, a common use of RNAi is in genome-wide screens using shRNA libraries, the goal being 70% knockdown of each target mRNA from each shRNA. Approximately 30% of shRNAs using a H1-promoter/RNAPIII perfectly-stemmed library attain the 70% level of knockdown (Bernards et al. 2006). In addition to the need for increases in efficacy, off-target effects due to heterogeneous processing are an area that needs optimization, otherwise producing false positives. Strong candidate genes, identified via siRNA to impact the pathway of a researcher's interest are not consistently replicated in pooled shRNA library screens, and the reverse is also true. Genes reported to be relevant to a pathway of interest via siRNA, but that cannot be replicated via shRNA, could merely be due to inefficient processing of the shRNA. However, genes reported to be

relevant via shRNA, but that cannot be replicated by siRNA, indicates inaccurate shRNA processing, and subsequent off-target effects may be the culprit. A strong indicator of this off-targeting is the low level of overlap between siRNA and shRNA screens; 80-95% of candidate genes identified by the pooled shRNA screen were not shared by the siRNA screen (Bhinder and Djaballah 2013; Bhinder et al. 2014a). One possible contribution to off-targeting, observed for a pLKO first generation shRNA: siRNAs replicating alternate processing sequences from the hairpin other than merely the guide strand reproduced the off-target knockdowns produced by the hairpin precursor (Bhinder et al. 2014b).

Beyond laboratory use, one of the most appealing uses for RNAi is antiviral gene therapy, particularly because RNAi is already the cell's innate immune response against viruses, and the promiscuous activity viewed as a negative elsewhere makes it more difficult for error-prone replication to allow antiviral escape. Research includes HIV-1 therapies, with an HCV therapy driven by an attenuated RNA polymerase III (RNAPIII) U6 promoter currently in clinical trials (Brake et al. 2006; Berkhout and Liu 2014; Borel et al. 2014; Lavender et al. 2012; Haussecker and Kay 2015). Deep-sequencing of the shRNA small RNA products from the HCV therapeutic was published in early 2014, showing a high level of heterogeneity: a range of 3-4 nucleotides for 5' and 3' strand positioning, inclusion of loop nucleotides that should have been removed by Dicer, and poor strand bias (i.e. both passenger and guide strands were produced). This high level of heterogeneity mimics that of the unintended guide strands, the siRNAs that replicated pLKO off-target effects (Bhinder et al. 2014b).

In the subsequent sections, we will discuss the toxicities and sources of toxicities associated with some previous attempts to harness RNAi. We will be comparing shRNAs to

shRNAmirs, and conclude with a discussion of AgoshRNAs due to our observations in

Chapters 2 and 3 of intermediates that overlap with the mechanism of AgoshRNA

biogenesis. shRNAs are generally stronger than shRNAmirs in terms of efficacy, but the

latter have closed much of the gap, while the toxicity, inefficient processing, and off-target

effects of the former remain (Boudreau et al. 2008a). Regardless, shRNAs have taught us a

lot about which portions of the biogenesis pathway can be toxically saturated, what can go

wrong with exogenous use of endogenous RNAi machinery. Though shRNAmirs are much

less toxic, through reduced levels of transcription and more endogenous-like structures, they

require more effort on the part of researchers to achieve the same levels of efficacy.

shRNAs and lessons from siRNAs

RNAi was discovered as dsRNA, sliced up to produce siRNA. A key step in

harnessing RNAi to silence genes was in producing siRNA from DNA, the main method

being linking a 5P strand to a 3P strand via a <10 nucleotide loop, producing short hairpin

RNA (shRNA) precursors, usually transcribed by RNAPIII. An early shRNA precursor had

19bp of uninterrupted RNA duplex stem, a 4 nt loop, and a 3' oligouridine termination

sequence 4-5 nts in length, and was designed to skip Drosha processing. Of particular note,

shRNAs were less effective than siRNAs at lower concentrations, an initial observation that

processing bottlenecks may exist (McManus et al. 2002; Brummelkamp et al. 2002; Paddison

et al. 2002). The RNAPIII-transcribed pLKO.1 shRNA library, maintained by The RNAi

Consortium, otherwise known as the Mission library, and the source for vectors used in

Chapter 3, is similar in structure to that previous described, consisting of a 21-mer stem, a 6nt loop, and the same 4-5 nt oligouridine terminus (Moffat et al. 2006).

The main goal of shRNA usage is to produce a guide strand that is incorporated into Ago-RISC. We learned a lot about this mechanism of incorporation and maturation of RISC, specific to Ago2, from siRNAs. While miRNAs are unwound, the passenger strand of an siRNA duplex is perfectly basepaired, and so can be cleaved, opposite nucleotides 9/10 from the 5' phosphate of the guide strand, accelerating the removal of the passenger strand. Cleavage-resistant bulged substrates, however, are rate-limiting because they must be unwound down a different, slower pathway (Matranga et al. 2005; Miyoshi et al. 2005; Rand et al. 2005; Rivas et al. 2005; Kim et al. 2007). This Ago2-specific cleavage phenomenon is further exacerbated in Ago1/3/4: increasing siRNA thermodynamic stability blocks passenger strand unwinding, while remaining sensitive to Ago2 cleavage (Petri et al. 2011).

Now, let us discuss some of the many ways shRNA expression can perturb something other than that which is intended. Exportin-5 was found to be a saturable portion of the miRNA/shRNA biogenesis pathway in the presence of high doses of AAV-introduced RNAPIII-transcribed shRNAs. This shRNA saturation of XPO-5 harkens back to Zeng and Cullen's 2004 interpretations of Exportin-5 as a protective reservoir for endogenous precursors, which when absent (or in the case of RNAPIII-transcribed shRNAs, already at capacity) might indicate global reductions in mature miRNA quantity. Subsequently, Ago2 was also found to be saturable; overexpression of both Ago2 and XPO-5 increased knockdown capability and stability and relieved toxicity. The secondary structure of RNA components also affected toxicity: high levels of guide:passenger complementarity, as well as guide:target complementarity, were toxic. This structure/toxicity relationship may be

connected to the Ago2 vs. Ago1/3/4 difference in terms of slicing ability (Ago2), perhaps the inability of Ago1/3/4 to unwind highly thermodynamic base-pairing (siRNAs and shRNAs) as sources of toxicity (Grimm et al. 2006, 2010).

Examples of shRNA-mediated toxicity abound: shRNA expression from an RNAPIII U6 promoter was found to contribute to MYC-induced carcinoma, globally disrupting miRNA expression (Beer et al. 2009). Hepatitis B virus treatment with adeno-associated virus gene therapy carrying U6-driven shRNAs was toxic, and reductions in transcription levels via the weaker RNAPII H1 promoter reduced toxicity, but with concurrent reductions in shRNA potency (Sun et al. 2013). *In utero* electroporation of shRNA, but not miRNA or siRNA produced off-target effects, disrupted endogenous miRNA levels, and suppressed let-7 production via saturated Dicer, which was alleviated via transfer of the guide/passenger strands to an endogenous miRNA loop + flank backbone (Baek et al. 2014). Yet another U6-driven report of toxicity was due to increased accumulation of the shRNA guide strand over 3 weeks (Ahn et al. 2011).

We have seen that reduced levels of transcription could avoid toxicity, but potency from the traditional shRNA structure in addition to lower levels of transcription is difficult: 1 of 380 tested worked in one study, when attempting to transition from a U6 to weaker H1 RNAPIII promoter. shRNA design parameters are not aligned for potency in conjunction with use at lower levels of expression (Shimizu et al. 2009). In a different report, the transition from U6 to H1 resulted in similar target knockdown, indicating a saturable bottleneck after which additional knockdown cannot be attained, regardless of transcription levels of this type of shRNA (Brake et al. 2008). Additionally, the RNAPIII promoters typically used for shRNA transcription are promiscuous in their position of initial nucleotide,

19

a range of 3-4 nucleotides for the U6 and H1 promoters, and an even larger range of initial

nucleotide position was observed for the modified/attenuated U6 promoter currently in anti-

HCV clinical trials (Ma et al. 2014a; Denise et al. 2014).

RNAPIII-transcribed shRNAs are effective, but likely due to brute-force transcription

of precursors, not through efficient processing. A researcher still has to test 3-4 to get a good

knockdown, with off-target effects that may not be immediately obvious. We have learned

that much of the endogenous biogenesis machinery can be saturated, displacing,

dysregulating, and globally repressing endogenous miRNA levels. All in all, traditional

shRNA design and usage parameters just do not seem to be aligned regarding high levels of

target repression, in conjunction with use at lower levels of shRNA expression. Next, we will

discuss the more efficient shRNAmirs, which have been optimized to compete with the high

level of knockdown initially produced from artificial, RNAPIII-transcribed shRNAs, but

without toxicity.

shRNAmirs: miRNA-like structures enhance tolerance and biogenesis

We will begin this section by discussing the main shRNAmir libraries commercially

available, then discuss the components of shRNAmir structure and processing that ablate the

toxicity and match the efficacy observed for shRNAs, and conclude this section with some

examples of miRNA-like stem structures that have previously been used to induce RNAi.

In comparison to shRNAs designed to bypass Drosha processing, shRNA-miRs (such

as the TRIPZ library based upon the miRNA-30 backbone) are meant to undergo all

endogenous miRNA processing steps. In the case of TRIPZ, the asymmetric bulge present in

endogenous miR-30 was not transferred to exogenous RNAi usage, but the flexible miR-30 loop and basal stem 5' and 3' Drosha flanking regions were, enhancing optimal Dicer and Drosha processing. This overall structure of perfect stem + flanks + loop has been found to trump traditional shRNA structures' potency, producing >12-fold more guide strand, particularly in the single-copy integrant context, and is commercially available as the TRIPZ/GIPZ library from OpenBiosystems (Zeng et al. 2002; Zhang and Zeng 2010; Zeng and Cullen 2003; Boden et al. 2004; Stegmeier et al. 2005; Chang et al. 2006; Silva et al. 2005).

An alternate miRNA backbone for shRNAs, miR-155/BIC allowed for intron-localization of shRNAmirs within a protein marker, as well as poly-cistronic production of up to eight tandem copies of shRNAs. However, in comparison to the miR-30 shRNAmir adaptation libraries, the stem structure of miR-155 was preserved for subsequent usage, maintaining a 5:3 asymmetric bulge in the center portion of the miRNA-shRNA stem for these 5P guide strands. These RNAPII-transcribed shRNAs are marketed, by Thermo Fisher Scientific under the BLOCK-iT brand, to be equally capable of repressing targets vs traditional RNAPIII-transcribed shRNAs (Chung et al. 2006).

Lower levels of transcription, miRNA loops, and Drosha-substrate flanks are all components of increased potency and/or reduced toxicity. Unlike shRNAs and siRNAs, shRNAmirs do not observably compete with endogenous miRNAs for transport and RISC incorporation (Castanotto et al. 2007). Dosed to produce the same level of RNAi activity, shRNAmirs are better tolerated vs. shRNAs, an example of the value of undergoing traditional processing events. Merely reducing transcription levels is not a cure-all: a weaker RNAPIII promoter, H1 (vs. U6), still promotes pre-miRNA build-up and toxicity *in vivo*

(Boudreau et al. 2008b). Maintaining the same high level of U6/RNAPIII transcription throughout the set of experiments, two active, toxic shRNAs maintained activity but lost toxicity when guide/passenger strands were embedded in a Drosha-substrate & miRNA loop context (McBride et al. 2008). Endogenous miRNA loop sequences by themselves increase processing efficiency (Hinton et al. 2008). The loop, bulged stems, and Drosha-substrate flanks in a polycistronic context, achieve for the shRNAmir the same level of knockdown as shRNA guide strands present at 15-30-fold higher levels (Liu et al. 2008).

Thermodynamically unstable stems, whether obtained via bulged/mismatched basepairs, or G:U wobble pairing, are miRNA-like structures; recommendations of their incorporation into shRNAs creates an shRNAmir. A random nucleotide library produced a structure with 3 mismatch bulges along its stem that were critical for activity (Wang et al. 2008a). In a different setting, mismatches at positions 1, 4, and 10, starting from the 5' phosphate, increased shRNA activity vs. perfectly stemmed precursors (Wu et al. 2011). Bulged stem shRNAmirs as part of bi-functional shRNAs, have been combined with traditional stem structure shRNAmirs, resulting in increases in potency, durability, and *in vivo* use without toxicity (Rao et al. 2010; Phadke et al. 2011).

Within the miR-30 backbone, though a recent report ended up recommending optimization of the Drosha substrate region for increased efficacy due to increased ease of adoption, the authors also found that the replacement of the endogenous asymmetric bulge back into miR-30 based shRNAmirs greatly increased RNAi activity (Fellmann et al. 2013). Elsewhere, a miR-30-based shRNA optimization paper wanted to take the backbone back to their roots, placing the guide strand back on the 5P strand, re-introducing the stem central symmetric bulge, and adding another bulge at position 4 paper, alterations that improved

RNAi activity (Zeng et al. 2013). These observations in addition to those in Chapter 2 produce an aggregate recommendation: return to miRNA-like stem structures, in addition to the loop and flank modifications, for optimal shRNAmir activity.

So why would thermodynamic instability be of positive benefit? Some of it has to do with enhanced processing and unwinding of the passenger from the guide strand, but there is also a contribution from Dicer-independent biogenesis that we will discuss in the next section. One perspective is that stem central bulges or G:U basepairs assist in Ago1/3/4 unwinding, the rate limiting step in RISC activation for these Ago proteins. With or without a single nucleotide central bulge in the stem, shRNAs associate with Ago1-4 indiscriminately, so Ago 1/3/4 might be clogged with perfectly stemmed RNA duplexes, resistant to unwinding and activation of RISC (Gu et al. 2011). It is interesting to note that this clogging was verified elsewhere with perfectly stemmed precursor saturating Ago3, indicating an Ago2 cleavage requirement for optimal RISC maturation of perfectly stemmed precursors (Yang et al. 2012).

We have learned a lot about how to take an shRNA and move it through the canonical endogenous miRNA processing pathway with as high processing efficiency and low toxicity as possible, primarily through the incorporation of Drosha substrate flanks and endogenous miRNA loops. Though mismatches in the stem have been recommended, their adoption by the community is far from overwhelming. Together, these developments in shRNAmir technology allow single-copy integration events to compete with the toxically saturating levels of RNAi activity produced by first generation shRNAs.

Dicer-Independent Biogenesis and AgoshRNAs

Previously, Ago2 was known to have the following activities: it bound the guide strand, it cleaved target mRNA, it repressed translation, it cleaved siRNA passenger strands for RISC maturation, if the stem central region was perfectly complementary, and it produced ac-cleaved pre-miRNA in very limited cases. Now we will learn of necessary non-canonical processing by Ago2 of an endogenous miRNA precursor. Though discovery of the non-canonical biogenesis pathway has produced a new mechanism by which to harness RNAi, it is also the likely culprit behind some of the inefficiencies that we report for shRNAmirs and shRNAs in Chapters 2 and 3.

In a surprising state of affairs, an endogenous microRNA that is processed by Drosha skips Dicer processing, instead maturing via Ago2. A precursor depositing complex surveys the precursor for Dicer processing compatibility (an abundance of stem mismatches induce the Dicer pathway) vs. potential mid-stem cleavage by Ago2 (perfectly complementary stems induce this pathway). Ago2 cleaves the 3P as if it were a passenger or target mRNA, and the terminal end is further trimmed to the loop by the poly(A)-specific ribonuclease, such that the 17nt 5' arm + 4nt loop combine to become the guide strand. This Dicer-independent miRNA precursor stands out amongst other endogenous microRNA precursors in that it looks similar to shRNAs in some aspects. Though shorter than an shRNA (17nt stem, too short for efficient Dicer processing), pre-miR-451 contains no mismatches in its stem. Introduction of a bulge at the site of Ago2 cleavage inhibits mature miR-451 production by Ago2. The miR-451 secondary structure can carry other sequences through the Dicer-independent pathway, much as the miR-30 backbone was intended to carry sequences through Dicer. Though processed at a lower efficiency compared with Dicer, potency may be

higher than Dicer-specific shRNAs due to Ago2-specificity (Cheloufi et al. 2010; Cifuentes et al. 2010; Yang et al. 2010; Dueck et al. 2012; Herrera-Carrillo et al. 2014, 2015; Liu et al. 2015a, 2013; Yoda et al. 2013; Liu et al. 2012) (Figure 1.2).
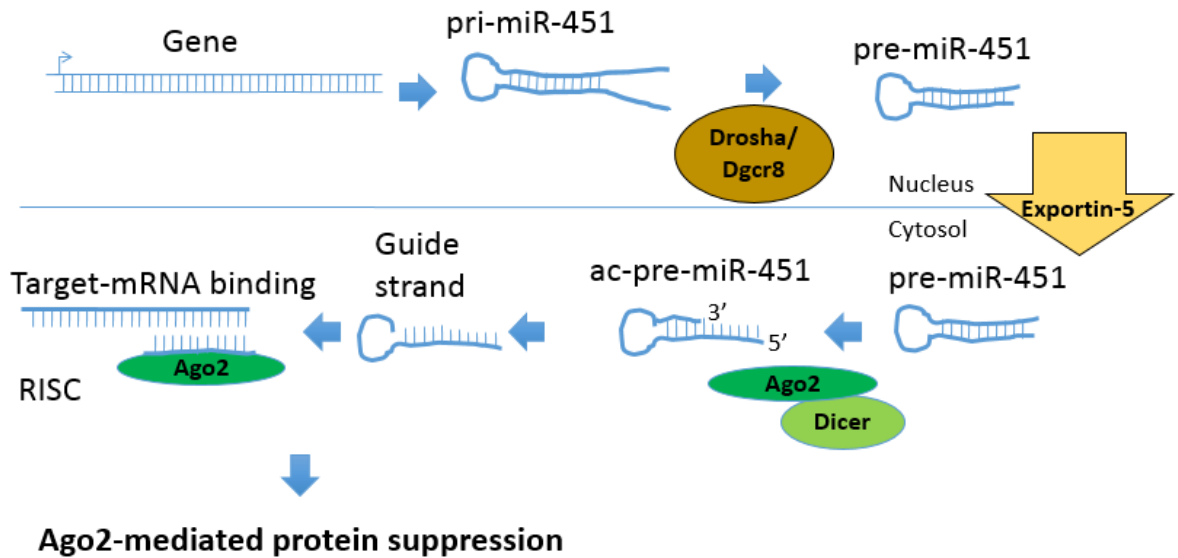
Figure 1.2 Dicer-independent miRNA biogenesis. Drosha and DGCR8 process the secondary structure as in the canonical pathway. The pre-miRNA transits the nuclear pore complex via Exportin-5, however Dicer-independent biogenesis requires a thermodynamically stem, optimally shorter so as to not compete with Dicer processing, i.e. ~17 basepairs in length. Short length + stability allow Ago2 to produce an ac-pre-miRNA, cleaving the 3P strand opposite nucleotides 10/11 of the 5P strand. The remainder of the 3P strand is recessed to the loop by PARN, after which the entirety of the 5P strand + the loop become the guide strand for Ago2. No other Ago proteins incorporate miR-451, because Ago2 is the only enzyme with slicing ability, and miR-451 is too short to transit the Dicer pathway.

AgoshRNAs and 5P guide strand short shRNAs (sshRNAs) are the different terms for shRNA variants that prefer the Dicer-independent pathway, but can be inefficiently processed by Dicer as well. Much as the guide strand of miR-451 is on the 5P, guide strands for sshRNAs are of increased potency if they arise from the 5P. Also similar to miR-451, mismatches in sshRNA stems reduces efficacy, but this effect does not inhibit efficacy if the stem is lengthened (to again mimic an shRNAmir and traverse the Dicer pathway). Additionally, placement of the guide strand on the 3P requires loop cleavage (Dicer maturation) for optimal activity, while 5P guide strands do not (Ago2 maturation) (Ge et al. 2010; Dallas et al. 2012).

miR-451 biogenesis can be disrupted through thermodynamic instability/central-stem mismatches. Though it is established that Dicer can inefficiently process AgoshRNAs, and Ago-2 can inefficiently process shRNAs, mainly due to stem length, negative effects of the Dicer-independent pathway on shRNAs and shRNAmirs have yet to be linked. A few hints exist; Dicer is definitely bypassed by some shRNAs, with RISC incorporation occurring in unintended ways (a portion of the 5' arm, the loop, a portion of the 3' arm, for example). A single RNAPIII TRC pLKO hairpin silenced six genes in addition to its target, due to inappropriate cleavage/maturation ending up in RISC, silencing targets (Bhinder et al. 2014b). In *Drosophila*, Dicer depletion results in guide strand variability increases including miR-451-like loci (Yang et al. 2014). Together, these observations point to the possibility that Dicer-independent biogenesis may affect shRNAs, not only inhibiting their own potency, but also dysregulating endogenous miRNA levels.

Chapter 2: Optimizing short hairpin RNA design to maximize RISC loading

Overview

Gene silencing via short hairpin mediated RNAi (shRNA) is a valuable experimental tool. One of the most popular shRNA platforms makes use of the loop and flanking sequences from the endogenous microRNA (miRNAs) miR-30a to provide an RNA structure for efficient biogenesis of the RNA trigger. However, the stem regions of these shRNAs are designed as perfect duplex structures which is an uncommon feature for endogenous miRNA precursors. Here, we report that shRNAs with perfect duplex stems undergo extensive stem cleavage analogous to the Dicer independent miRNA miR-451. This cleavage event destroys the shRNA trigger sequence that is present in the 3P arm. We employed an unbiased screen of shRNA structures to identify features that prevent stem cleavage and promote canonical biogenesis and loading into the effector complex RISC. We find that a central stem bulge in addition to a secondary bulge is an optimized structure for shRNA efficiency.

Introduction

RNA interference (RNAi) is a widely used technology to study gene function (Hannon and Rossi 2004; Mohr and Perrimon 2012; Kaelin 2012). Broadly, RNAi can be

divided into two experimental approaches based on the form of the gene silencing trigger. siRNAs are short duplex RNAs that are transiently transfected into the experimental cell line. After incorporation into the RNA induced silencing complex (RISC), the guide strand of the siRNA directs RISC to complementary target RNAs and promotes silencing of the target gene. This form of RNAi is relatively short lived due to the transient nature of the siRNA trigger. In contrast, plasmid encoded shRNA triggers are capable of long term gene silencing. shRNAs are stem-loop RNAs that require cellular biogenesis pathways prior to RISC loading. shRNA expression systems based on retroviral vectors are especially valuable since they can be used for silencing gene expression long-term in cells refractory to transient transfection.

shRNA expression vectors are based on two general designs. The original shRNA vectors used the U6 or H1 promoter to drive expression of a stem loop RNA with a 19 or 29 nucleotide stem and a small loop (Moffat et al. 2006; Bernards et al. 2006; Paddison et al. 2002; Brummelkamp et al. 2002). These RNAs resemble miRNA precursors and are directly processed by the miRNA biogenesis enzyme Dicer prior to RISC loading. More recent vector designs have the shRNA trigger embedded in the loop and flanking sequences of miR-30a (shRNAmir) (Zeng et al. 2002; Stegmeier et al. 2005; Silva et al. 2005). This shRNA transcript resembles a miRNA primary transcript and requires Drosha and Dicer processing prior to RISC loading. shRNAmirs are typically expressed from a RNA polymerase II promoter and therefore allow the use of the full complement of regulatable mammalian promoter systems. Additionally, the lower transcription rate of RNA polymerase II promoters and more natural loop and flanking regions result in reduced toxicity of the RNA compared to U6 based shRNA vector systems (Sun et al. 2013; Ahn et al. 2011; Shimizu et

29

al. 2009; Boudreau et al. 2008a; McBride et al. 2008; Castanotto et al. 2007). Notably, both vector designs frequently place the shRNA trigger sequence on the bottom of the stem loop (3P strand) for more accurate and biogenesis cleavage products. Furthermore, both shRNA designs commonly maintain perfectly duplexed stem structure unlike endogenous miRNAs (including miR-30a). While retroviral shRNA systems have found widespread use in biomedical research they often display incomplete target gene knockdown, especially when integrated at a single genomic copy (Fellmann et al. 2011, 2013; Premsrirut et al. 2011).

We have previously developed next generation sequencing technologies for the study of miRNA biogenesis (Newman et al. 2011). Here, we applied these approaches to the analysis of shRNA biogenesis. We find that the vast majority of shRNA precursors are internally cut on the 3P stem analogous to the Dicer-independent miRNA miR-451. Since shRNA vector systems are designed with the gene silencing trigger on the 3P strand, this cleavage will limit the amount of trigger RNA that can be loaded into RISC. To improve shRNA biogenesis efficiency we developed an unbiased screen to identify structural features that promote Dicer processing and RISC loading. We report that central mismatches are essential to prevent internal 3P cleavage and that an additional mismatch improves shRNA yield and knockdown potential.

Materials and Methods

Cell culture, transfection, and transduction

3T3 fibroblasts were cultured as described by ATCC.org. Cells were transduced with TRIPZ as described by Dharmacon, using 1:3 serial dilutions into 6-well plates to obtain single-copy MOI. Cells were selected and maintained in 1.25 ug/ml puromycin. 1×10^5 cells were plated in 24-well format on Day 1 of a reporter assay; shRNAmir expression was induced after cell attachment with 1 ug/ml doxycycline (replenished every 24 hours until lysate harvest). On Day 2, psiCHECK 2.0 constructs were transfected using Lipofectamine 2000. On Day 3, media was changed, and the lysate was harvested on Day 4 using Promega Passive Lysis Buffer. For co-transfection studies, 1ug/ml doxycycline was present during transfection of both the TRIPZ and psiCHECK 2.0 plasmids, otherwise usage was identical to Day 3 onward of the transduction protocol. All luciferase assays within a single panel of a figure compare single-copy integrants of the same passage number (i.e. control, parental, and experimental shRNAs were transduced side by side), and exposed to doxycycline and puromycin premixed prior to use into a common source. Error bars are the standard deviation of technical replicates of three independent transfections of reporter.

Constructs

Appropriate perfectly complementary 1X targets were designed to anneal and produce sticky ends ready for psiCHECK 2.0 XhoI/NotI. Guide and passenger strand sequences were obtained from the RNAi Consortium, pLKO.1-based plasmids TRCN0000090506 (Col1A1) and TRCN0000254936 (Fox3). These sequences were transferred to the miR-30/TRIPZ backbone, amplified with 5' and 3' primers containing

31

EcoRI/XhoI sites, and ligated into TRIPZ. To maintain desired secondary structure while still producing identical 3' guide strands throughout, sequence variants that transferred responsibility for this structure to the 5' passenger strand were verified by RNAfold (Gruber et al. 2008; Lorenz et al. 2011). All shRNAmir and reporter oligonucleotides used can be found in S1 Table.

High-throughput deep sequencing and analysis

Precursor-specific deep sequencing was performed as previously described (Newman et al. 2011). Ago-associated RNA were immunopurified using Protein A Dynabeads (Life Technologies) and the 2A8 antibody, preferential for Ago2 (Diagenode). RNA was extracted using Trizol (Invitrogen). Small RNA deep sequencing libraries originated from 5.0 ug total RNA using a method similar to that previously published (Pfeffer et al. 2005). Briefly, a 3' adapter was ligated to total RNA using truncated RNA ligase 2, which was then size extracted (retaining 40-60nt products of linker + small RNA and removing excess 3' adapter) on a 12.5% 8M urea polyacrylamide gel. T4 RNA ligase 1 attached a 5' adapter containing a two nt barcode. After reverse transcription and PCR amplification, these barcodes allowed multiplexing prior to sequencing on a HiSeq 2000 (Illumina). Primers that amplified the flanking regions of shRNAs were used to verify plasmid library screen depth, as well as to identify and quantify genomic integrants in the population for normalization; see S1 Table.

Raw sequence reads were processed using a custom bioinformatics pipeline (Newman et al. 2011). Precursor reads were mapped according to terminal nucleotide position. Mature shRNAs from the library screen were mapped to the originating structural variant. Reads for each variant were normalized to the amount of integrated provirus for that variant.

Results

       We investigated the production of mature shRNA triggers from the widely used

TRIPZ retroviral vector. This vector has the shRNA duplex sequence within the context of

the miR-30a loop and flanking sequences. Transcription is driven from a tetracycline

inducible promoter; see plasmid map in Figure 2.1. We studied two moderately weak

shRNAs that target the Col1A1 and Fox3 mRNAs. Single copy integration into NIH-3T3

cells led to 51 and 43% knockdown, respectively, for the two shRNAs (Figure 2.2A). As

expected, introduction of additional copies of the shRNA vector by transient transfection

improved target gene knockdown. We prepared small RNA libraries for next generation

sequencing to quantify the amount of mature shRNA trigger present in stably expressing

cells. We observed 1521 and 185 reads per million for the Col1A1 and Fox3 shRNAs,

respectively, in cells with single copy integration of the corresponding TRIPZ vector. These

reads are in contrast to endogenous miRNAs that in some cases had 100 fold more mature

sequence reads. Notably, miR-21a is a miRNA present at a single genomic locus yet had

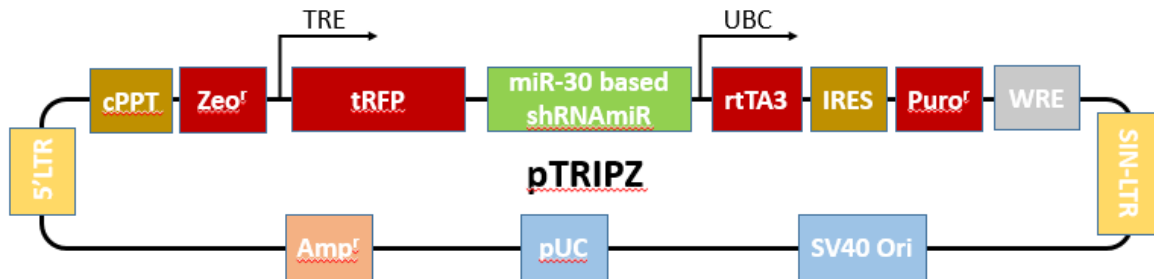greater than 200,000 normalized sequence reads.

Figure 2.1. pTRIPZ plasmid map. 5'LTR and SIN-LTR are long terminal repeats; SIN-LTR is self-inactivating. The Tetracycline-response element contains 5 tet operators and a minimal CMV promoter; the miR-30 based shRNAmir is embedded in the 3' UTR of turbo RFP. Bacteria are cultured in Carbenecillin and Zeocin; cells are cultured in Puromycin. (Adapted from Dharmacon Technical Manual)

We next sought to identify causes for the low yield of mature shRNA trigger in virus transduced cells. We employed a previously developed deep sequencing approach specific for miRNA precursors. This approach provides not only read counts for specific precursor sequences but also provides information about internal 3P stem cleavage, trimming, and nontemplated nucleotide addition. For example, analysis of precursor reads for the miRNA Let-7g demonstrates that almost all precursors are full length sequences, with < 1 percent of the sequences trimmed and/or cut internally in the 3P stem or near the loop (Figure 2.2C, blue bars). Let-7a1 has a greater fraction of precursors that are cleaved internally; however, most precursors remain full length. The increase in internal cleavage for Let-7a1 is likely due to the perfect duplex structure of the stem-loop, a feature known to promote Ago2 mediated cleavage.
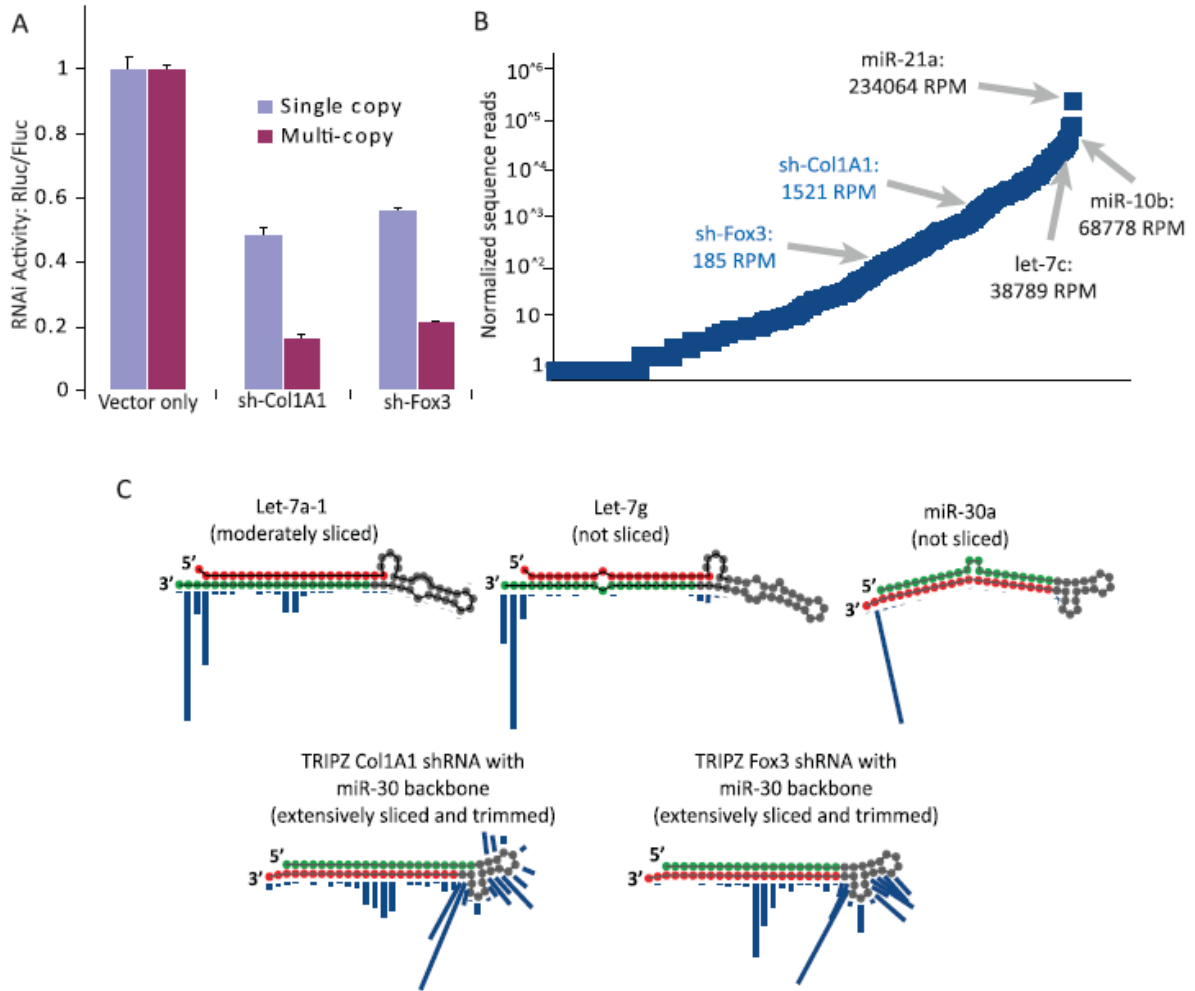
Figure 2.2. Stem cleavage reduces biogenesis efficiency of shRNAs. **(A)** Knockdown activity for two different shRNAs was determined using a luciferase reporter vector. NIH-3T3 cells were stably transduced with shRNA retrovirus at single copy integration and target knockdown measured. Multi-copy knockdown was determined by transient transfection of additional shRNA vector into cells. **(B)** Small RNAs from shRNA transduced cells were sequenced using the Illumina platform. Normalized miRNA and shRNA reads are plotted in rank order. **(C)** Precursor miRNAs and shRNAs were sequenced using a customized Illumina protocol. Read counts for individual miRNAs/shRNAs were counted for each terminal nucleotide position on the precursor sequence. The height of each blue bar indicates the relative read count for precursors that terminate at the indicated 3' position. Three representative miRNAs are shown along with the shRNAs.

35

We then analyzed precursor reads for the two shRNAs. Over 99 percent of reads had evidence of internal cleavage and trimming. This cleavage pattern is reminiscent of the Dicer independent miR-451 biogenesis and may be related to a recently reported Ago2-based precursor surveillance system (Cheloufi et al. 2010; Cifuentes et al. 2010; Yang et al. 2010; Liu et al. 2012, 2014). Notably, the endogenous precursor for miR-30a, which the TRIPZ system is based on, contains an asymmetric bulge at nucleotides 12/13 of the 5P arm and is resistant to internal cleavage. However, when the miR-30a precursor was adapted to the shRNA vector system the central stem bulge was not retained; only the loop and flanking sequences were retained. The stem region of these shRNAs are perfect duplexes and therefore substrates of internal Ago2 cleavage, a critical problem, since the trigger strand is the 3P strand of the stem. Any internal cleavage will destroy the trigger, reducing potency and potentially increasing incorporation of the 5P strand with concomitant off-target effects. It should be noted that trigger strand cleavage is not only observed with TRIPZ. We find extensive cleavage with the U6 based pLKO vector - which also has a perfect duplex stem (data not shown).

To directly test whether the perfect duplex stem was promoting internal cleavage we modified the shRNA vector to include a central bulge in the duplex region in the range of positions 9-11, positions that inhibit Ago2 cleavage of siRNA passenger strands, target mRNAs, and miR-451 biogenesis (Martinez et al. 2002; Matranga et al. 2005; Elbashir et al. 2001; Cifuentes et al. 2010; Yang et al. 2010; Leuschner et al. 2006). We introduced these mismatches in the 5P non-trigger strand of both the Col1A1 and Fox3 shRNA vectors. The 3P trigger sequence was unchanged to allow direct comparison to the original structure. Precursor sequencing of these modified shRNAs demonstrated that > 99 percent of reads

were full length with very limited internal cleavage (Figure 2.3A). We directly compared

knockdown efficiency of the stem-bulged shRNAs with perfect duplex shRNAs. At single

copy integration, bulged shRNAs led to increased knockdown potency (Figure 2.3B).
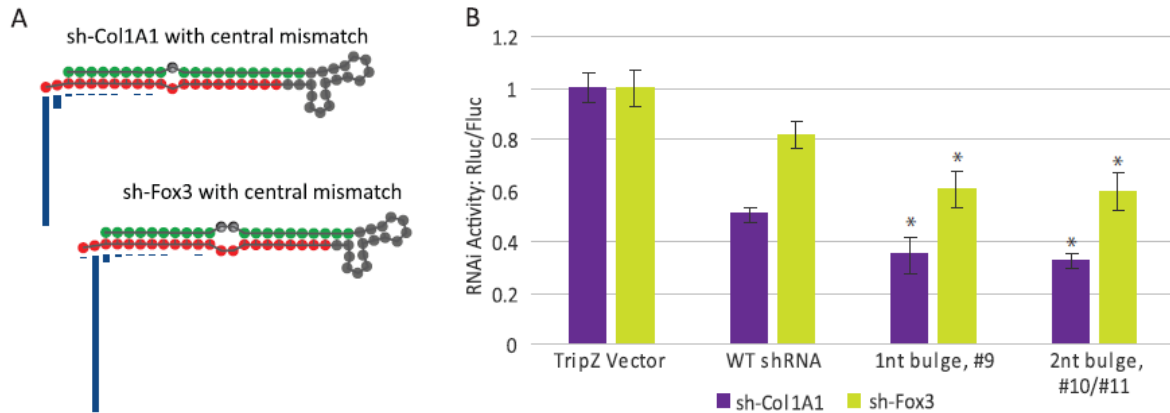
Figure 2.3 shRNA stem-central bulges prevent internal cleavage and promote knockdown activity. (A) Mutations were introduced into precursor sequences to create a central bulge in the stem loop structure. A single-nucleotide bulge at #9, counting from the 5' terminal end, and two nucleotide bulges at #10/11 were introduced in the 5P arm to allow identical trigger sequences. Precursor shRNAs were sequenced and read counts for individual shRNAs were counted for each terminal nucleotide position on the precursor sequence. The height of each blue bar indicates the relative read count for precursors that terminate at the indicated 3' position. **(B)** Knockdown activity was determined of mutant shRNAs structures from (A). shRNAs were transduced at single copy and reporter activity measured. Asterisk indicates statistical significance of difference in activity compared to WT shRNA ($P < .05$, Mann-Whitney-U test).

While these rational design changes improved shRNA performance, we wanted to identify other structural features that would further optimize shRNA biogenesis and RISC loading. To this end we devised a non-biased screen to identify optimal shRNA structures (Fig 2.4A). Through the use of degenerate nucleotides, we generated a library of different shRNA structures around the Col1A1 and Fox3 shRNA sequences within the TRIPZ vector. These structures included central bulge options at stem positions 9-11 counting from the 5' precursor phosphate. Since all three positions were degenerate, the library contains non-bulged structures as well as single, double, and triple mismatch-containing bulges. We also wanted to determine whether additional miRNA precursor-like characteristics would increase shRNA trigger loading. Therefore, libraries included a degenerate nucleotide at position 19, allowing a single nucleotide bulge at the Dicer cleavage site. Libraries also included degenerate nucleotides allowing an additional bulge 3' and 5' of the central bulge. These additional bulges could be mismatches (MM), G:U base pairs, or could be WT (non-bulged). All degenerate nucleotides were introduced on the 3P trigger strand; the 5P strand was unaltered for all structures (Fig 2.4B). We introduced this library of 144 possible structures for each shRNA into NIH-3T3 cells by retroviral transduction. To identify structural variants that enhanced biogenesis and RISC loading, we isolated Ago2 associated RNAs by immunoprecipitation, and identified enriched degenerate nucleotide structures, normalized to the sequenced provirus population.
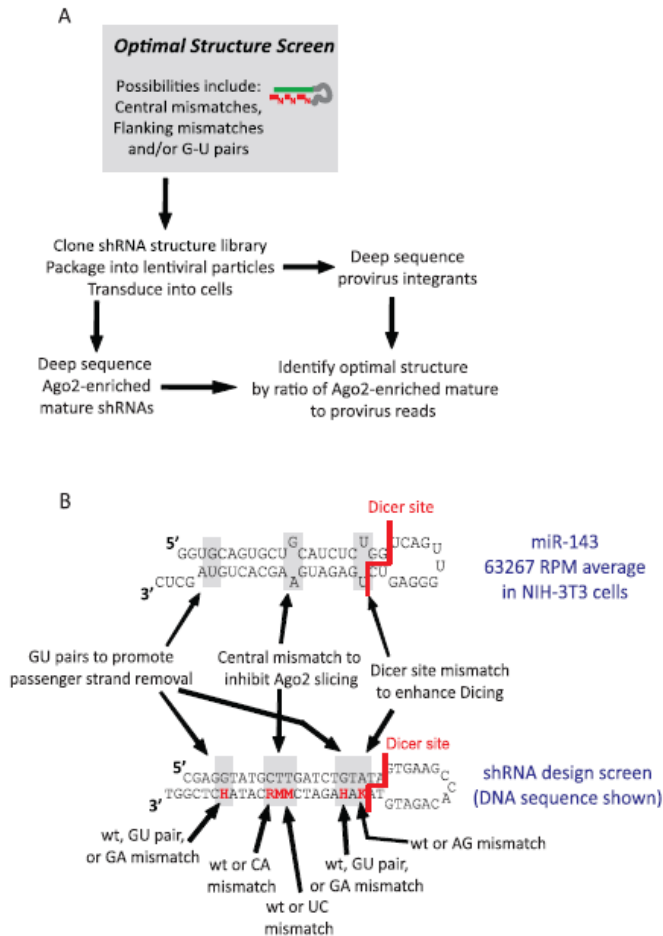
Figure 2.4. Unbiased screen for optimized shRNA structures. (A) The flow chart for our shRNA structure screen is shown. Randomized shRNA structures were generated using degenerate oligonucleotides and were cloned into the TRIPZ vector backbone. Structure libraries were packaged into virus and transduced into NIH-3T3 cells. Ago2 associated small RNAs were isolated and identified by deep sequencing. Optimal structures were identified by normalizing deep sequencing reads of Ago-associated guide strands to provirus reads. **(B)** The structural variants in the library are illustrated and compared to an abundant miRNA precursor.

The most striking feature of enriched shRNAs was a central bulged sequence. Figure 2.5A displays fold enrichment of shRNA structures containing specific central bulges. Each bar represents the average enrichment of the 18 structures in the library that contain that specific central bulge (3x3x2 structures for 5' and 3' flanking bulge/GU and Dicer bulge). Therefore, some individual structures with the indicated central bulge have a greater enrichment than shown since some structures were heavily disfavored. Also worth nothing within this panel, the WT bars in the graph (non-bulged) show enrichment > 1 due to the favorable effects of flanking bulges increasing the average for the entire bin. A direct comparison of each central bulge position when all other degenerate positions were wild type is shown in the supplement (Figure 2.S1A).
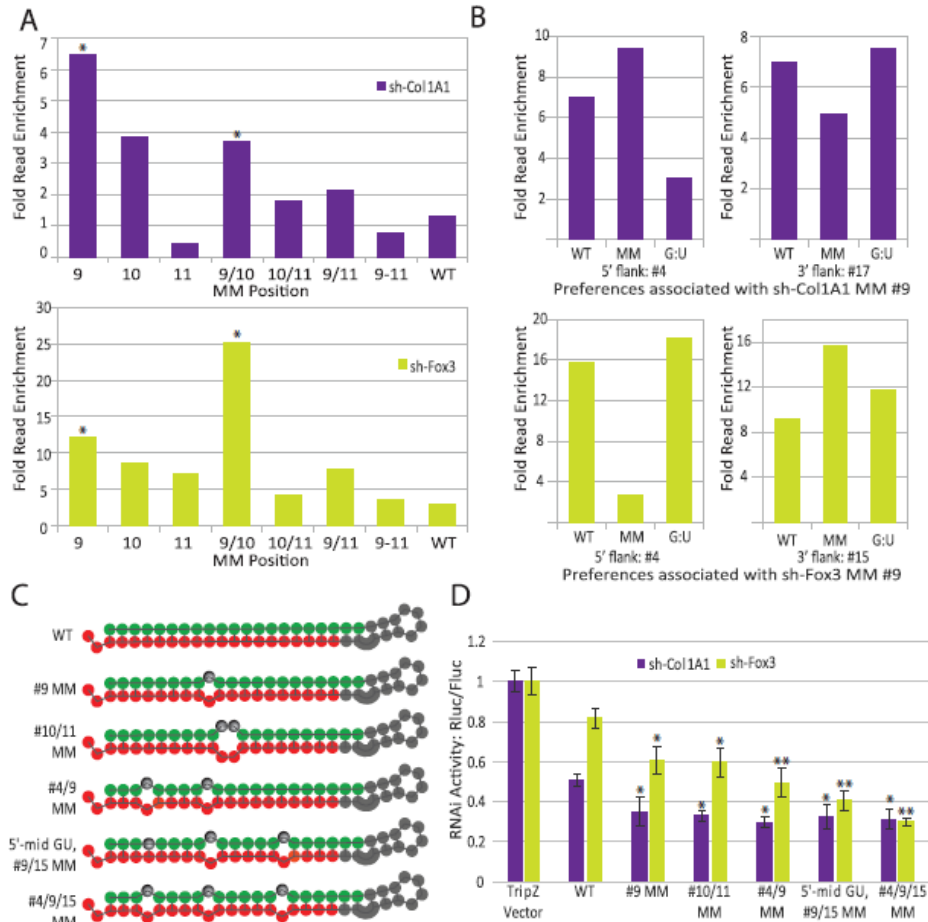
Figure 2.5. Optimized shRNA structures and activity. (**A**) Ago-associated shRNAs were sequenced to determine the optimal structures for biogenesis and RISC loading. Structural variants were binned by type of central mismatch and average fold enrichment versus full wild type sequence is shown. Asterisk indicates significance compared to WT central stem bulge, $P < .05$, student's paired two-tailed t-test. (**B**) Structural variants with a mismatch at nucleotide position #9 were further binned by 5' and 3' flanking mismatch variant. Fold enrichment of read counts versus full wild type is shown. (**C**) Illustration of precursor structures that are presented in the figure. (**D**) Knockdown efficiency of enriched structural variants. shRNA structures were cloned by moving the indicated mismatches into the 5P stem, allowing identical trigger sequences. shRNAs were transduced into NIH-3T3 cells at single copy and knockdown determined by reporter gene activity. A single asterisk indicates statistical significance of difference in activity compared to WT shRNA; two asterisks indicate significance compared to both WT shRNA and shRNAs with only central-stem bulges ($P < .05$, Mann-Whitney-U test).

41

Immunoprecipitation of Ago-associated reads revealed a preference for a bulge for the Col1A1 and Fox3 shRNAs at nucleotides 9 and 10, though the Col1A1 shRNA was optimal with a single nucleotide bulge at 9, while the Fox3 shRNA was optimal with a double nucleotide bulge at 9/10. Adding a flanking bulge provided modest further enrichment in Ago2 associated shRNAs, though the two shRNAs differed in the preferred location of the flanking second bulge. The Col1A1 shRNA preferred a second bulge 5' to the central bulge while the Fox3 shRNA preferred a second bulge on the 3' flank in addition to a 5' G:U. Disparities here likely have to do with the relative thermodynamic stabilities of each shRNA sequence's 5' end competing with the benefits of inhibited slicing activity, but overall agree with structural determinants for endogenous miRNA RISC loading and slicer-independent unwinding (Schwarz et al. 2003; Kawamata et al. 2009).

Structures are illustrated in Figure 2.5C. Bulges very close to the Dicer cleavage site were strongly disfavored (not shown), supported by previous reports of Dicer promiscuity induced when the intended cleavage site is within 2 nucleotide of a loop or bulge (Gu et al. 2012).

shRNA structural improvements were validated by reporter knockdown analysis. The enriched shRNA structures were recapitulated by switching the mismatch to the 5P arm to allow a common WT 3P trigger sequence. Both the Col1A1 and Fox3 shRNA vectors were modified to contain the same structural alterations. Retroviral vectors were transduced into NIH-3T3 cells at single copy MOI and knockdown determined by reporter gene expression. Structural features that lead to increased enrichment by Ago2 associated deep sequencing also provided improved knockdown (Figure 2.5D). Notably, both shRNAs were improved

by the same structural alterations even though their preferences as reported by the screen were slightly different.

Discussion

In this report we describe extensive internal cleavage and trimming of shRNA precursors. The cleavage is likely due to direct endonucleolytic action of Ago2 on the shRNA precursor. The Dicer independent miRNA miR-451 is processed by Ago2; several miRNAs have been reported to be cut in the 3P stem at some frequency (Cheloufi et al. 2010; Cifuentes et al. 2010; Yang et al. 2010; Diederichs and Haber 2007). Interestingly, incompletely trimmed miR-451 (up to 30nt in length after Ago2 cleavage), retains RNAi activity (Yoda et al. 2013). Regardless of how much of the remaining non-canonically processed precursor maintains activity, many have used this discovery to circumvent Dicer processing, using the miR-451 structure as a template for alternatively processed shRNAs (AgoshRNAs) carrying 5P+loop guide strands (Ge et al. 2010; Dallas et al. 2012; Liu et al. 2013; Ma et al. 2014b; Herrera-Carrillo et al. 2014; Yang et al. 2012; Liu et al. 2015a). Ago2 cleavage of precursors is dependent on a perfect duplex stem, a feature that is absent in most endogenous miRNA precursors. In the case of miR-451, Ago2 cleavage of the 3P stem can lead to productive mature miRNA, since the mature strand is on the 5P stem. The most common shRNA vector designs, however, place the active shRNA trigger on the 3P stem. Thus, Ago2 cleavage would destroy the trigger and compromise the potency of the shRNA. While the structural features that promote Ago2 cleavage of precursors has been known, our

deep sequencing results described herein demonstrate the extent of precursor loss and resultant inefficiency of conventional shRNA designs.

Several groups have used rational design to improve shRNA structures for increased target gene knockdown. Stem mismatches have been previously reported to increase knockdown potency (Wang et al. 2008b; Wu et al. 2011; Zeng et al. 2013; Rao et al. 2010). The stem bulges have been hypothesized to decrease Ago2 cleavage and/or to decrease thermodynamic pairing energy. Asymmetric mutations in the stem leads to a kinked and bulged structure that reduces accuracy of Dicer cleavage (Starega-Roslan et al. 2011). G:U wobble basepairs have also been described in stems to promote Dicer processing, may have the same effect as bulges when it comes to guide:passenger strand selection, and have been reported to inhibit AgoshRNA activity (Okamura et al. 2009; Gu et al. 2011; Liu et al. 2015b). While these improvements have been demonstrated, the adoption of these design changes to commercial shRNA vector systems has been slow. The miR-30-based TRIPZ vector as well as U6 based vectors (pLKO from the RNAi Consortium) are still designed with perfect duplex stems. One reason for the resistance to new structures is that shRNA libraries would need to be redeveloped. Furthermore, it is not clear that previous studies were able to maximize the design improvement due to the rational method of design changes. The unbiased structure library screen that we report here uncovers an optimized structure based on RISC loading enrichment.

In addition to increases in knockdown potency, reduction in Ago2 precursor cleavage may reduce unwanted off-target effects (Bhinder et al. 2014b). Small RNAs processed in unintended ways from a traditional shRNA currently in clinical trials parallel the sequences we observe in precursor-specific deep sequencing; surely an antisense drug would be

preferable if it only traversed one biogenesis pathway instead of two (Denise et al. 2014). Ago2/RISC activity has been reported to arise directly from pre-let-7a-3 without Dicer presence (Tan et al. 2009, 2011). Interestingly, Dicer depletion in Drosophila increased levels of Dicer-independent processing of miRNA precursors (Yang et al. 2014). If unintended Ago2 cleavage and subsequent RISC incorporation of shRNAs contributes to off-target effects, expression of shRNAs may be exacerbating these effects through Dicer saturation. It has been previously established that shRNAs can compete with miRNAs for export to the cytoplasm, and that shRNA toxicity can arise via saturation of the endogenous miRNA biogenesis pathway (Grimm et al. 2006, 2010; Boudreau et al. 2008a). It is interesting to speculate that saturation of Ago2 with perfectly duplexed shRNA may contribute in novel ways to shRNA toxicity, off-target effects, and/or miRNA pathway alterations.
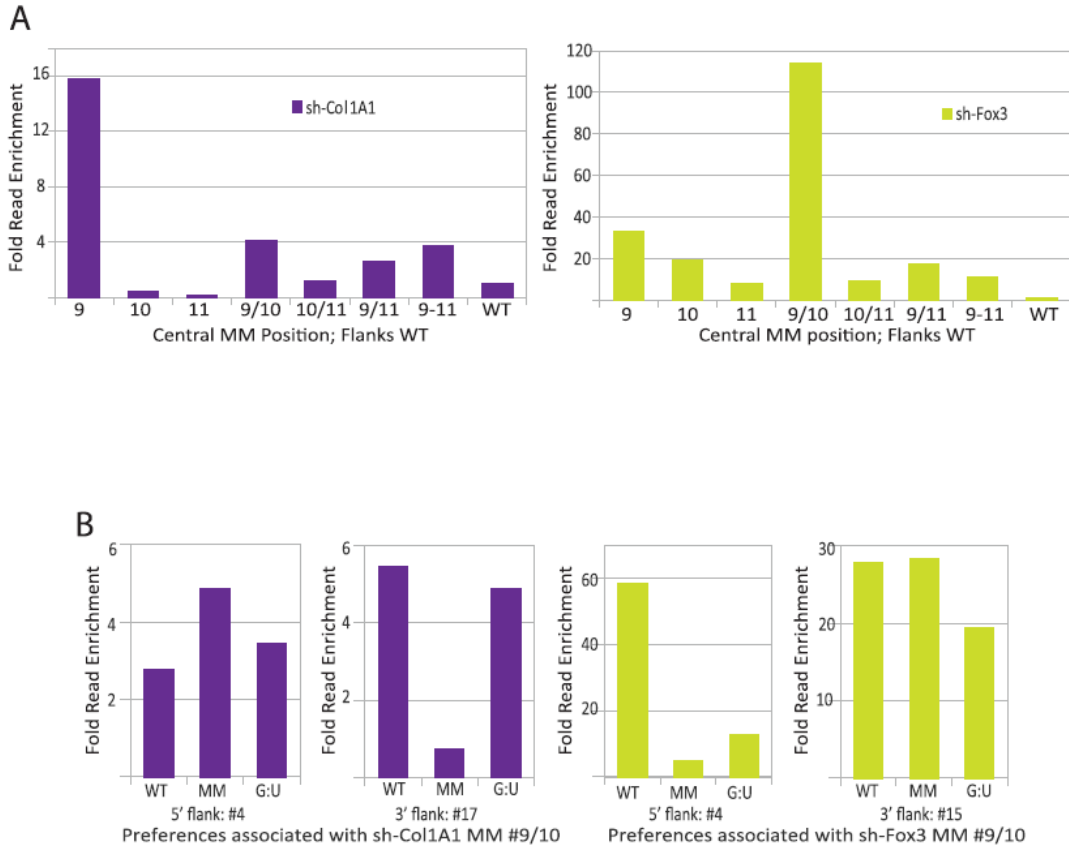
Figure. 2.S1 Additional structural preferences identified in screen. (A) Optimal structural variants, identified by enrichment of Ago-associated reads in comparison to provirus reads; bins here are all 5' and 3' WT flanks, such that the only variation is in the central-stem mismatch, and the WT bin would therefore be identical to the traditional TRIPZ structure. **(B)** Structural variants with a mismatch at nucleotide positions #10/11 were further binned by 5' and 3' flanking mismatch variant. Fold enrichment of read counts versus flanking wild type is shown.

Table 2.1: Cloning Oligonucleotides

All primers listed in 5' to 3' orientation. Italic = miR30 loop. Bold = miR30 flanking sequence. Lower case = non-WT nucleotide.	
Amplification primers	
5'miR30-XhoI	TACAATA**CTCGAGAAGGTATATTGCTGTTGACAGTGAGCG**
3'miR30-EcoRTI	ACTTAGAA**GAATTCCGAGGCAGTAGGCA**
Perfect and centrally-bulged stems	
Col1A1 WT	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGGTATGCTTGATCTGTA*TAGTGAAGCCACAGATGTA*TACAGATCAAGCATACCTCGG**TGCCTACTGCCTCGGAATTC**
Col1A1, bulge #9	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGGTATGaTTGATCTGTA*TAGTGAAGCCACAGATGTA*TACAG
Col1A1, bulge #10/11	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGGTATGCccGATCTGTA*TAGTGAAGCCACAGATGTA*TACAG
Col1A1 rev/comp 3' strand	**GAATTCCGAGGCAGTAGGCA**CCGAGGTATGCTTGATCTGTA*TACATCTGTGGCTTCACTA*T
Fox3 WT	**GGTATATTGCTGTTGACAGTGAGCG**GGATAGGTGGAGTAGGGTTAA*TAGTGAAGCCACAGATGTA*TTAACCCTACTCCACCTATCC**TGCCTACTGCCTCGGAATTC**
Fox3, bulge #9	**GGTATATTGCTGTTGACAGTGAGCG**GGATAGGTGGcGTAGGGTTAA*TAGTGAAGCCACAGATGTA*TTAACC
Fox3, bulge #10/11	**GGTATATTGCTGTTGACAGTGAGCG**GGATAGGTGGAagAGGGTTAA*TAGTGAAGCCACAGATGTA*TTAACC
Fox3 rev/comp 3' strand	**GAATTCCGAGGCAGTAGGCA**GGATAGGTGGAGTAGGGTTAA*TACATCTGTGGCTTCACTA*
Screen:	
Col1A1 Screen 5'+lp	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGGTATGCTTGATCTGTA*TAGTGAAGCCACAGATGTA*
Col1A1 Screen 3'+lp	**GAATTCCGAGGCAGTAGGCA**CCGAGdTATGykkGATCTdTm*TACATCTGTGGCTTCACTA*TACA
Fox3 Screen 5'+lp	**GGTATATTGCTGTTGACAGTGAGCG**GGATAGGTGGAGTAGGGTTAA*TAGTGAAGCCACAGATGTA*
Fox3 Screen 3'+lp	**GAATTCCGAGGCAGTAGGCA**GGATAdGTGGmkkAGGdTTAm*TACATCTGTGGCTTCACTA*TTAAC
Move screen preferences to 5' strand, amplify using Col- and Fox-specific rev/comp 3' oligos above.	
Col1A1, bulge#4/9	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGaTATGaTTGATCTGTA*TAGTGAAGCCACAGATGTA*TACAG
Col1A1, G:U #2, bulge #9/15	**GGTATATTGCTGTTGACAGTGAGCG**CCGgGGTATGaTTGATaTGTA*TAGTGAAGCCACAGATGTA*TACAG

Col1A1, bulge #4/9/15	**GGTATATTGCTGTTGACAGTGAGCG**CCGAGaTATGaTTGATaTGTA*TAGTGAAGCCACAGATGTA*TACAG
Fox3, bulge #4/9	**GGTATATTGCTGTTGACAGTGAGCG**GGATAaGTGGcGTAGGGTTAA*TAGTGAAGCCACAGATGTA*TTAACC
Fox3, G:U #3, bulge #9/15	**GGTATATTGCTGTTGACAGTGAGCG**GGATgGGTGGcGTAGGaTTAA*TAGTGAAGCCACAGATGTA*TTAACC
Fox3, bulge #4/9/15	**GGTATATTGCTGTTGACAGTGAGCG**GGATAaGTGGcGTAGGaTTAA*TAGTGAAGCCACAGATGTA*TTAACC
Psicheck 2.0	
Col1A1 for/rev	TCGAGCCGAGGTATGCTTGATCTGTA, GGCCTACAGATCAAGCATACCTCGGC
Fox3 for/rev	TCGAGGGATAGGTGGAGTAGGGTTAA, GGCCTTAACCCTACTCCACCTATCCC
DEEP SEQUENCING OLIGONUCLEOTIDES (in addition to those from Newman 2011)	
shRNA precursor-specific:	
Col1A1 WT	CAAGCAGAAGACGGCATACGACCGAGGTATGCTTGATCTGTA
Fox3 WT	CAAGCAGAAGACGGCATACGAGGATAGGTGGAGTAGGGTTAA
Col1a, bulge #9	CAAGCAGAAGACGGCATACGACCGAGGTATG**A**TTGATCTGTA
Fox3, bulge #10/11	CAAGCAGAAGACGGCATACGAGGATAGGTGGA**AG**AGGGTTAA
Provirus normalization (nn = barcode):	
PCR2 + nn + 3' miR30	AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGACGATCnnCTTGAATTCCGAGGCAGTAGGC
PCR1 + 5' mir30 sense	caagcagaagacggcatacgacttgctgggattacttcttcagg

Table 2.2. Chapter 2 summary
Pre-shRNA degradation inhibits RISC loading and activity
miR-30 based shRNAs are extensively degraded, reminiscent of Ago2-like cleavage and trimming.   **Mid-stem bulges known to inhibit Ago2-cleavage of 3P strand rescue precursor degradation and increase shRNA potency.**   **Unbiased design screen reveals RISC-loading preference for miR-30 based shRNA stem structures with weaker thermodynamics:**   ○ **Strong preference for mid-stem bulge**   ○ **Weaker preference for additional base-pairing disturbance flanking the mid-stem bulge.**

Chapter 3: Inefficiencies Associated with RNAPIII shRNA Design

Overview

Traditional shRNA structures are still in wide use, both for individual and pooled

library knockdown gene function experiments, in addition to preclinical and clinical trials.

Though we observe the same degradation of traditional shRNA precursors as we do for miR-

30 based shRNAs, RNAi activity is already strong due to U6-driven RNAPIII transcription,

and unaffected by degradation-inhibiting mid-stem bulges. Additionally, we find that the

oligouridine tail, which terminates transcription of a traditional shRNA, is incorporated into

guide strands, inhibits processing of precursors, and induces precursor degradation. Finally,

we observe dysregulation of the precursor surveillance system in the presence of traditional

shRNA expression: endogenous precursor oligouridylation levels increase and trimming

decreases. Overall, these data extend the depth of knowledge surrounding traditional shRNA

processing, as well as downstream negative consequences associated with endogenous

miRNA precursor surveillance.

Introduction

Traditional shRNA-mediated RNAi was a hop and a skip in tool development away

from siRNAs, ~21 nucleotide RNA duplexes that bypasses upstream processing machinery to

directly interact with Ago2 upon transfection, after which the intended inactive passenger

strands were cleaved and dissociated. However, transfection is time-limited, and not compatible with all cell types (Wang, 2009). Traditional shRNAs were engineered by connecting the 5P to 3P siRNA RNA strands via a 4-6 nucleotide loop allowed transcription by cell machinery, extending the amount of time available to maintain knockdown, and expanding potential knockdown environments in tissue culture and *in vivo* via transduction of cells previously refractory to transfection. However, this tool is only active after multiple biogenesis steps between introduction of the shRNA and activity, steps during which shRNA competes with endogenous RNAi species, promoting inefficient biogenesis of both, and in some cases inducing toxicity. Because the endogenous RNAi machinery is occupied, endogenous mRNAs normally repressed by endogenous miRNAs can be released in the presence of shRNA, high levels of precursor shRNAs can be transcribed but unprocessed, overloading the capacity of Exportin-5, Dicer, and the Ago proteins, provoking toxicity (Grimm et al. 2006, 2010; Beer et al. 2009).

These traditional shRNA expression vectors are based on two general designs, both without bulges, and still in use today, even though bulged precursors can produce similar levels of knockdown via Agos1-4, while perfectly stemmed precursors are limited to optimal activity in Ago1/2 (Paddison et al. 2002; Brummelkamp et al. 2002; Su et al. 2009). Three recent papers reveal the efficiencies of traditional shRNA design: heterogeneity of transcription initiation positioning, Dicer-independent unintended positioning of processed products, and off-target effects due to this unintended processing (Bhinder et al. 2014b; Denise et al. 2014; Ma et al. 2014a). Alternatively, Dicer maturation can be circumvented via the miR-451 biogenesis pathway, limiting the guide strand activity to Ago2 (Liu et al. 2015b; Herrera-Carrillo et al. 2014, 2015; Liu et al. 2015a; Yang et al. 2010, 2012; Liu et al. 2013).

Transcription of traditional shRNAs terminates at a five thymidine termination sequence, generating a shortened transcript intended to skip the first step of endogenous miRNA biogenesis, Drosha cleavage. Of interest is the transcript's oligouridine tail, a substrate for DIS3L exonuclease activity discovered in the post-transcriptional repression of mature let-7 processing (Ustianenko et al. 2013). An oligouridine tail of ~20 nucleotides in length was also previously reported to inhibit Dicer processing of let-7; what impact would this oligouridine tail have on processing (Heo et al. 2009)? Additionally, a recently reported precursor surveillance system, separate from the Lin28/let-7 regulatory pathway, checks for appropriate 3'-OH overhang length. Exosome component RRP6 degrades shorter overhangs, while DIS3 degrades longer overhangs that can be added by TUT-4/7, repressing guide strand expression levels (Liu et al. 2014).

Here, we use precursor-specific deep sequencing as previously mentioned to observe the same high levels of slicing and trimming of traditional shRNAs' 3P arm encoding the RNAi trigger. Likely due the saturating transcription rates of U6-driven RNAPIII that produce toxicity, high levels of knockdown are observed with and without the addition of a mid-stem bulge. We also observe heterogeneity in guide strand positioning, in agreement with previous reports, a trait of these traditional shRNAs that seems to be ignored. *In vitro* introduction of uridyl-tailed precursors show other negative details associated with traditional shRNA design: inhibition of Dicer processing and accelerated degradation that align with DIS3-associated precursor surveillance. This precursor surveillance system provides an interesting explanation for our final observation: that the mere expression of these traditional shRNA structures perturbs endogenous surveillance activity, inducing increases in 3' oligouridylation and decreases in trimming of endogenous miRNA precursors. These data

52

together contribute to a recommendation to depart from traditional shRNA structures in research and clinical use, instead embracing more miRNA-like shRNAs, shRNAs that don't straddle Dicer-independent vs. Dicer-dependent pathways.

Materials and Methods

Deep sequencing and luciferase assays were completed as described in Chapter 2, as well as in a previous publication from our lab (Newman et al. 2011).

In vitro transcription time course

In vitro transcripts for precursors with the appropriate length uridine tails were produced as described previously (Newman et al. 2008). Equivalent amounts of gel-purified, internally-radiolabelled precursors were exposed to S100 lysates from N1E, HeLa, or P19 cell lines, or immuno-precipitated Dicer, as indicated. At the noted time points, equivalent fractions of a master mix were spiked into Trizol, precipitated, and run on a denaturing polyacrylamide gel, dried, and imaged. Graphs of degradation or Dicer products were produced via quantification of bands using ImageJ software. Error bars are the standard deviations from independent experiments.

In vitro transcripts were produced with annealing of a forward binding primer to the full-length oligo. Sequence for T7 forward primer: gactagTAATACGACTCACTATA. WT sequences are shown; 3' uridines were added as labeled for each precursor species. mmu-miR-103-1: 5'-TAATACGACTCACTATAGGCTTCTTTACAGTGCTGCCTTGTTGCAT ATGGATCAAGCAGCATTGTACAGGGCTATGA-3'. Mmu-miR-27b: 5'-TAATACGAC TCACTATAAGAGCTTAGCTGATTGGTGAACAGTGATTGGTTTCCGCTTTGTTCAC

AGTGGCTAAGTTCTGC-3'. Mmu-miR-25: 5'-TAATACGACTCACTATAAGGCGGAG

ACTTGGGCAATTGCTGGACGCTGCCCTGGGCATTGCACTTGTCTCGGTCTGA-3'.

Results

We studied the same guide/passenger strands as in Chapter 2, but within the pLKO,

U6-RNAPIII traditional shRNA context, again observing high levels of cleavage and

trimming, with >99% of precursors' trigger strands degraded. This is likely Ago-associated,

not due to the 5U tail, as a similar length sequence lacking secondary structure but

maintaining the 5U tail was relatively stable in its length (non-templated additions not

included in this positional analysis) (Fig. 3.1A). We observed high levels of guide strand

production, but heterogeneity in guide strand positioning, with ~99% of 5' initial nucleotide

of guide strand beginning one nucleotide later than expected. Heterogeneity of positioning at

3' terminus of guide strand is almost entirely templated but non-targeting 3' oligo-uridine.

Together, these changes produce guide strands of ~24 nts in length, as opposed to the

expected 21. Figure 3.1B presents the combined precursor-specific deep sequencing position

of 3' terminus as a percentage of total reads for both Col1A1 and Fox3 shRNAs, combined,

in comparison to the oligo-uridyl tailed, unstructured transcript (Fig 3.1B). As expected due

to the high levels of guide strand produced (data not shown), a bulged stem replaced the

perfectly base-paired stem without changing the high level of reporter activity (Fig. 3.1C).
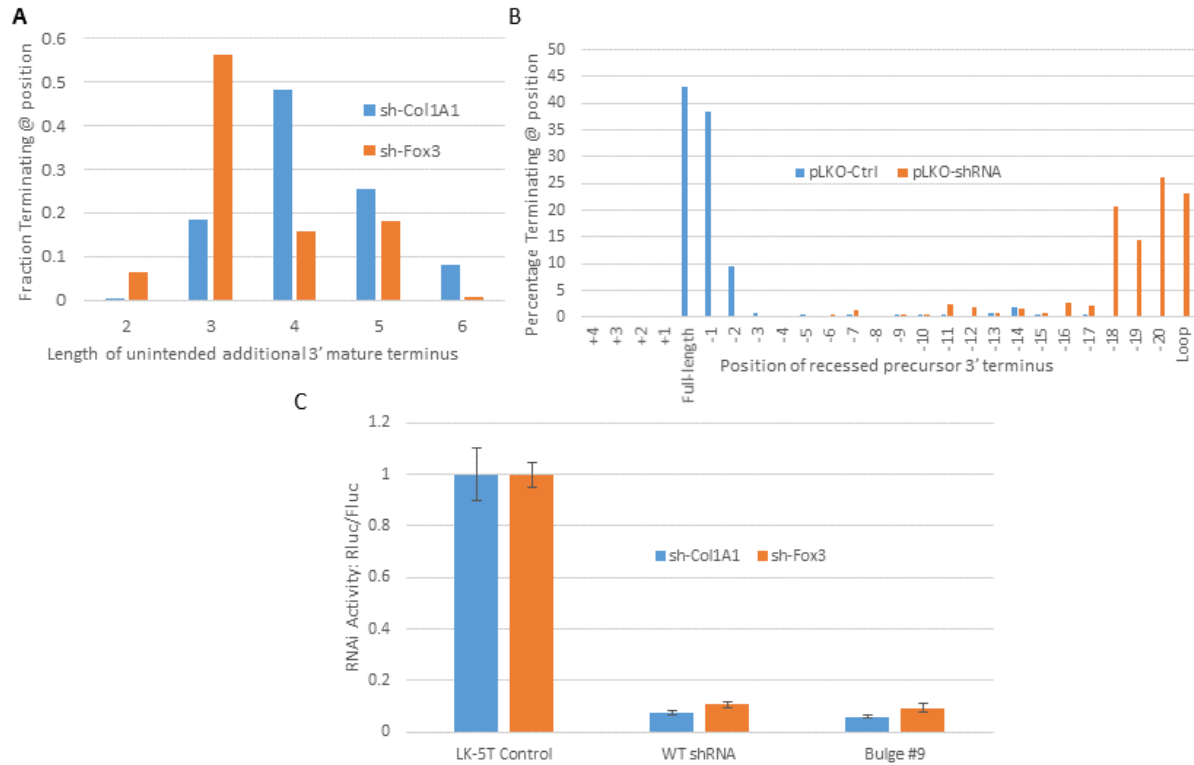
Figure 3.1: High levels of RNAi activity undeterred by heterogeneous processing. (A)
Both Col1A1 and Fox3 shRNA precursor deep sequencing reads were binned by 3' terminus
position as a position of their total reads. Positional percentages were added and then halved
to present a combined snapshot of positional data for perfectly-stemmed shRNA precursors.
This was compared to the positional preferences as a percentage of that species' total for
unstructured, empty-vector, 5-uridine tailed transcript. **(B)** Position of 5' and 3' terminus of
intended 3P guide strands for sh-Col1A1 and sh-Fox3, is shown as a fraction of that species'
total reads. **(C)** Single-copy transduced 3T3 cells, transfected with PsiCHECK reporter as in
Chapter 2, comparing perfectly stemmed Col1A1 and Fox3 reporter activity to single
nucleotide bulge at position #9 from the 5' phosphate, normalized to unstructured pLKO-5T
empty vector.

To isolate the contribution of the uridine tail to degradation and Dicer processing, we utilized *in vitro* transcription of uridine-tailed pre-miR-25 and pre-miR-27b, exposed to S100 crude cell extracts for 20 minute time courses, revealing decreases in stability inversely proportional to the length of the tail (Fig 3.2A). Tails of 3-uridines and longer were compared to wild type precursors without untemplated additions, and tails of all sizes inhibited the production of mature-sized products. Additionally, these same time courses showed reductions in production of mature-sized miRNAs. To isolate whether this inhibition of Dicer processing wasn't due to degradation, we exposed the radiolabeled tailed and untailed precursors to immunoprecipitated Dicer, observing a direct inhibition on processing by the tail (Fig 3.2B). Of particular note, a 3-uridine tail, added to an endogenous precursor's WT 2 nucleotide overhang, shown to inhibit Dicer processing, produces the total length of overhang as encoded by traditional RNAPIII shRNAs.
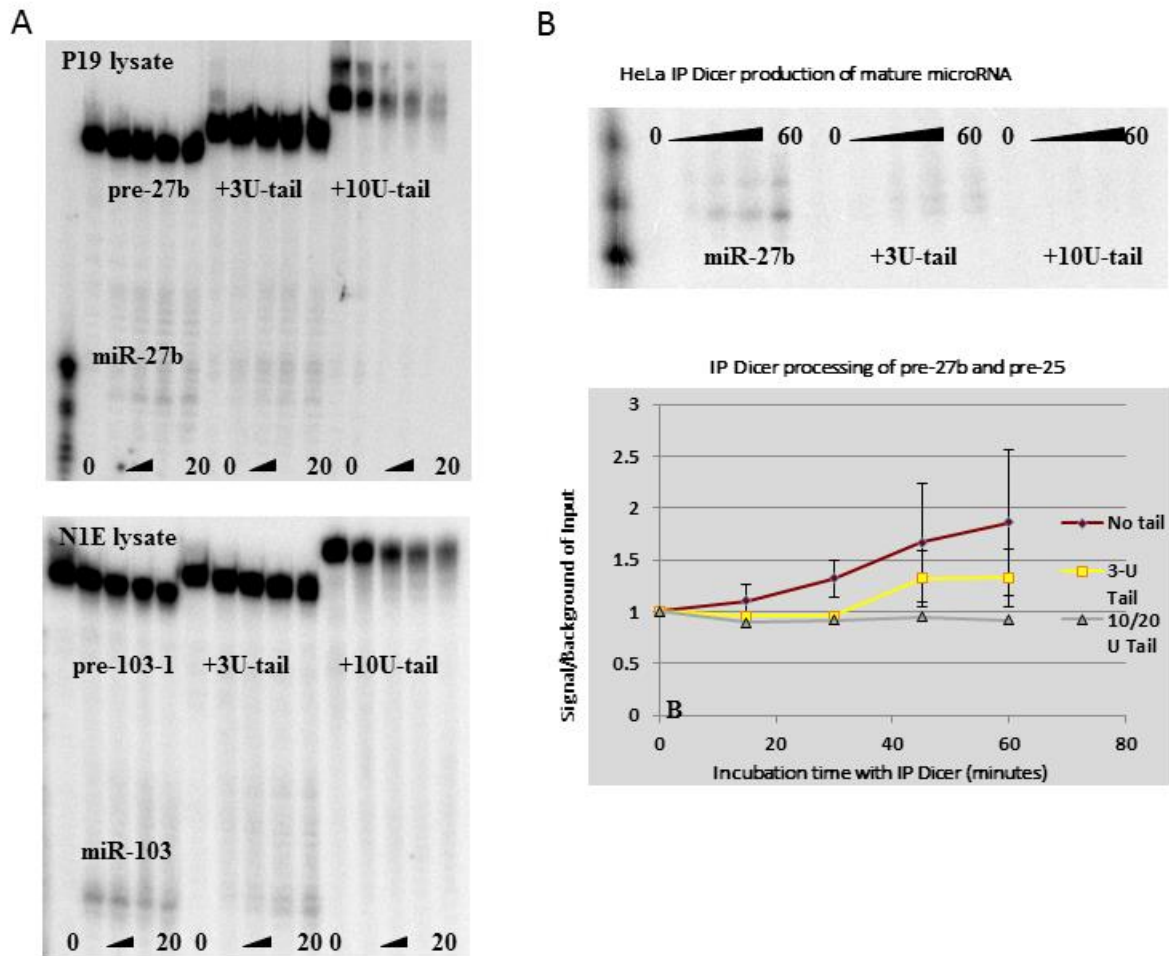
Figure 3.2: Uridyl-tailed precursors induce degradation and inhibit Dicer processing *in vitro*. (A) Radiolabeled, *in vitro* transcribed pre-miR-27b and -25 with uridine tails of labeled lengths were exposed to N1E, HeLa, or P19 crude S100 cell lysates, with the leftmost lane for an initial lane time point 0, and subsequent lanes dividing the remaining time evenly. **(B)** Dicer was immunoprecipitated from HeLa lysates to isolate processing from degradatory activity. A 60-minute time course gel exposure with pre-miR-27b is shown; the graph combines ImageJ quantitation of Dicer product-sized from pre-miR-27b and pre-miR-25.

A large fraction of endogenous miRNA precursors terminate in a tri-uridine tail, a feature that we have shown to negatively impact stability, and separately, Dicer processing *in vitro*, which aligns with the DIS3 exonuclease preference for >2nt overhangs. In a TUT-4 knockdown, we observed changes in the tailing profile: total precursors reads that were oligouridylated minimally decreased from 5% to 3.5% for the TUT-4 shRNA-expressing cells, compared to WT (Fig 3.3A). Regardless, deep sequencing supported our *in vitro* time course observations that oligouridylation decreased stability. The precursors with the most dynamic changes in oligouridylation after TUT-4 shRNA expression show an inverse correlation with total reads (Fig 3.3A). However, the startling observations came from non-targeting (scramble and anti-GFP) shRNAs: we observed an increase in overall endogenous miRNA precursor oligouridylation vs. WT cells. Upon further analysis, though a higher percentage of all endogenous precursors were oligouridylated in shRNA-expressing cells, trimming was reduced (Fig 3.3B). Subsequently, we ceased attempting to use shRNAs to knockdown RNAi uridylation machinery, and verified via different non-targeting shRNAs, perturbation of precursor miRNA surveillance: oligo-uridylation is induced, while 3'-5' degradation of endogenous precursors is inhibited (Fig 3.3C).
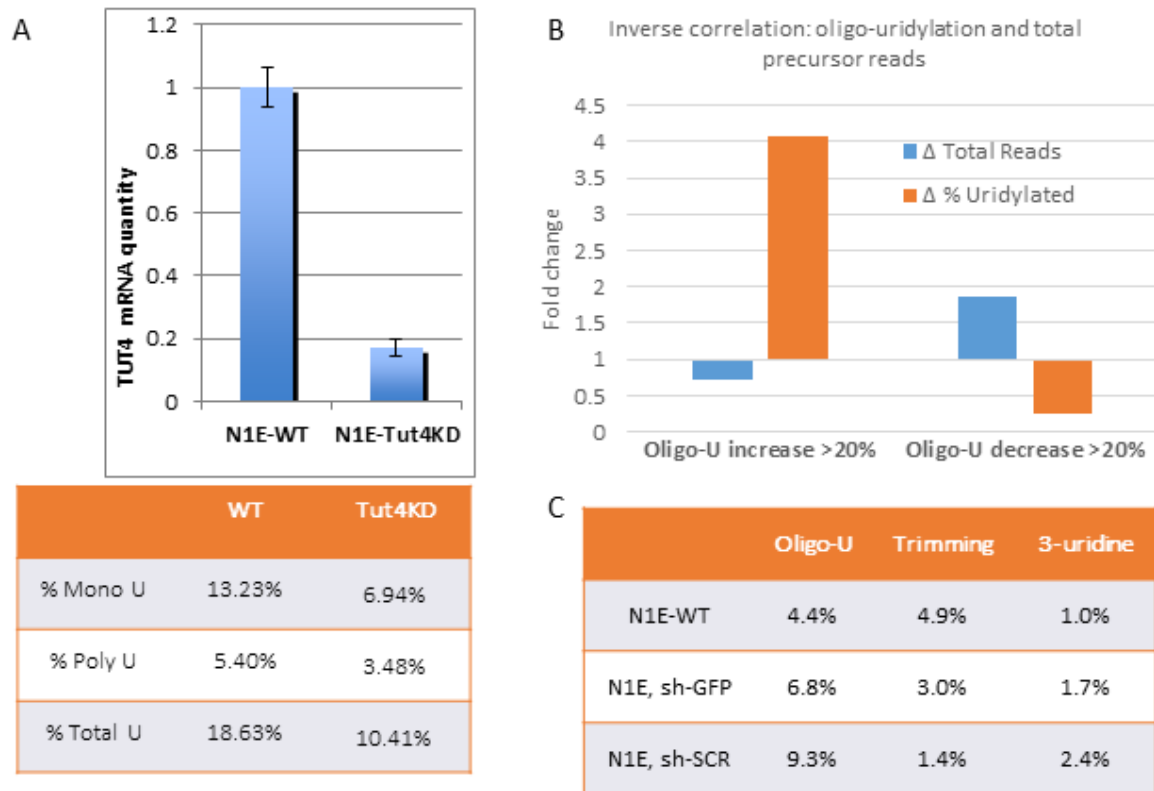
A

	WT	Tut4KD
% Mono U	13.23%	6.94%
% Poly U	5.40%	3.48%
% Total U	18.63%	10.41%

B Inverse correlation: oligo-uridylation and total precursor reads

C

	Oligo-U	Trimming	3-uridine
N1E-WT	4.4%	4.9%	1.0%
N1E, sh-GFP	6.8%	3.0%	1.7%
N1E, sh-SCR	9.3%	1.4%	2.4%

Figure 3.3: shRNA expression perturbs endogenous precursor miRNAs' tailing and degradation. (A) qPCR-verified knockdown of TUT-4 mRNA in N1E cells stably transduced with pLKO sh-TUT4; RNA from which was used to produce a precursor-specific deep sequencing library, with the total reads, and percentage of total that are uridylated shown. **(B)** Pre-miRNAs with dynamic changes in oligouridylation after sh-TUT4 expression in N1E cells were segregated into two bins: those that increased oligouridylation by more than 20% (17 of 219 pre-miRNAs tracked) vs. those that decreased oligouridylation by more than 20% (41 of 219). Fold change in fraction oligouridylated is shown for each bin, in comparison to fold change in total reads per million, for each bin. **(C)** Percentage of total endogenous pre-miRNA reads for oligo-uridylation, trimming, and tri-uridylation, are shown for the N1E cells expressing the labeled shRNA.

Discussion

Using precursor-specific deep sequencing, we report extensive internal cleavage and trimming of perfectly stemmed U6/RNAPIII shRNAs, independent of the degradative 5-uridine tail signal. Furthermore, shRNA expression perturbs endogenous precursor uridylation and trimming. These oligouridylated precursors are no longer optimal Dicer substrates; and should undergo DIS3 3'-5' degradation. If they are processed by Dicer and are 3P miRNAs, targeting activity will be modified due to the non-templated uridine tail, heterogeneity that occurs for 3P traditional shRNAs as well.

It's interesting to speculate whether Ago1/3/4 are clogged with precursor shRNA due to their difficulty unwinding highly stable duplexes, reducing endogenous miRNA activity. It would be very interesting to test total small RNA vs. Ago-associated RNA quantities in the presence or absence of traditional shRNAs. Recently, another report of RNAi pathway defects due to shRNA expression was released: dysregulated endo-siRNA expression and spermatogenic defects; they didn't observe any changes in miRNA expression levels via qPCR (Song et al. 2015). However, as previously published, and as we employed in Chapter 2, the Ago-associated fraction of small RNAs indicates potential activity, total quantity does not; endogenous miRNA activity may be reduced in the presence of shRNA expression, but not observed via qPCR (Flores et al. 2014).

If traditional shRNA design isn't modified, it's likely that Ago2 produces the unintended miR-451 like guide strands from the 5P + loop instead of the 3P, explaining the heterogeneity in the guide strands for the HCV therapeutic as well as the off-target effects arising from a different pLKO/TRC clone (Denise et al. 2014; Bhinder et al. 2014b). Off-target effects produced via non-canonical, Dicer-independent processing, or the

heterogeneity of the TT-034 HCV therapeutic, should be reduced via introduction of a mid-stem bulge that inhibits Ago-like slicing and subsequent trimming of the 3P, with no change in efficacy as we report at single-copy.

We observed high levels of knockdown for traditional shRNAs, leaving little room for improvements in increased RNAi activity, even at single-genomic copy contexts. However, an H1 promoter that drives RNAPIII is a popular alternative to U6, alleviating toxicity through reduced levels of transcription, but also losing efficacy, likely due through both decreased levels of transcription and increased heterogeneity in initiating nucleotide position of the transcript (Shimizu et al. 2009; Ma et al. 2014a). Alternately, the U6 promoter in the HCV therapeutic was modified to attenuate its level of transcription to alleviate toxicity (Lavender et al. 2012). Increases in RNAi efficacy observed after blocking degradation as in Chapter 2, via incorporation of mid-stem bulges, may still assist traditional shRNA efficacy if they're transcribed from weaker promoters than U6.

Perturbation of endogenous precursor modifications counterintuitive to the current model indicates saturation, perhaps of DIS3. If traditional shRNAs are competing for exosome activity, it would be interesting to determine whether substrates outside of RNAi are affected (Lee et al. 2014; Kiss and Andrulis 2010). Together, these data suggest that traditional shRNAs are a subpar tool with which to dissect the mechanisms and regulation of RNAi, and may also have unintended effects on other RNA biology pathways that use oligouridylation as a degradative signal/substrate.

One of the bullet points that ends up in the pro column when shRNAs are compared to shRNAmirs is the perceived accuracy and precision of cleavage positioning. It was previously assumed that skipping Drosha processing would allow the most effective,

accurate, and precise production of the desired guide strand. The 5U tail and heterogeneity observed with small RNA initiation of transcription produce guide strands with large variations between actual and expected products. In comparison, increased knowledge surrounding substrate preferences that produce accurate and precise Drosha and Dicer cleavage products, in addition to our recommended inclusion of mid-stem bulge to inhibit Ago-like degradation of the 3P strand, should allow a researcher to obtain much more precise and accurate *in vitro* and *in vivo* production of desired guide strands (Gu et al. 2012; Ma et al. 2013).

A recent paper recommended returning as close to the miR-30 backbone as possible, placing the trigger on the 5P strand, re-inserting a 5' bulge as well as a mid-stem asymmetric bulge (Zeng et al. 2013). This should inhibit unintended 3P cleavage that could allow incorporation of the loop into the guide strand via miR-451-like biogenesis, and the additional 5' bulge should enhance strand bias for the 5P as well (Schwarz et al. 2003). Now that we know more about Dicer and Drosha preferences, increasing the accuracy and precision of cleavages from engineered substrates, in combination with the heterogeneity arising from Dicer-dependent and –independent pathways presented here and in Chapter 2, future harnessing of RNAi should not entail usage of traditional U6/RNAPIII shRNAs.

Regardless of whether AgoshRNAs or Dicer-only shRNAmirs turn out to be optimal for research and/or therapeutic use, it seems unwise to use shRNA substrates that straddle both mechanisms. If a traditional shRNA is designed with a 3P trigger, cleavage destroys the 3P strand. If a traditional shRNA is designed with a 5P trigger, cleavage allows incorporation of the loop and potentially some of the 3P strand into RISC, modifying the RNAi activity profile.

Table 3.1. Chapter 3 summary
Traditional U6/RNAPIII shRNAs inefficiently harness RNAi
Traditional shRNAs are extensively degraded.   **Inclusion of degradation-inhibiting mid-stem bulges doesn't increase RNAi activity from the high level produced by traditional structure.**   **shRNA expression dysregulates precursor surveillance: oligouridylation of endogenous pre-miRNAs increases and degradation decreases.**   **Oligouridyl tail added to endogenous precursors and encoded into traditional shRNA 3' terminus induces clearance and inhibits Dicer processing.**

REFERENCES

Ahn M, Witting SR, Ruiz R, Saxena R, Morral N. 2011. Constitutive Expression of Short Hairpin RNA in Vivo Triggers Buildup of Mature Hairpin Molecules. *Hum Gene Ther* **22**: 1483–1497.

Altuvia Y. 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research* **33**: 2697–2706.

Ambrogio A D', Gu W, Udagawa T, Mello CC, Richter JD. 2012. Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. *Cell Reports* **2**: 1537–1545.

Ameres SL, Horwich MD, Hung J-H, Xu J, Ghildiyal M, Weng Z, Zamore PD. 2010. Target RNA-Directed Trimming and Tailing of Small Silencing RNAs. *Science* **328**: 1534–1539.

Aravin A, Tuschl T. 2005. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Letters* **579**: 5830–5840.

Auyeung VC, Ulitsky I, McGeary SE, Bartel DP. 2013. Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. *Cell* **152**: 844–858.

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* **455**: 64–71.

Baek ST, Kerjan G, Bielas SL, Lee JE, Fenstermaker AG, Novarino G, Gleeson JG. 2014. Off-target effect of doublecortin family shRNA on neuronal migration associated with endogenous microRNA dysregulation. *Neuron* **82**: 1255–1262.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.

Bartel DP. 2009. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136**: 215–233.

Basyuk E, Suavet F, Doglio A, Bordonné R, Bertrand E. 2003. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* **31**: 6593–6597.

Beer S, Bellovin DI, Lee JS, Komatsubara K, Wang LS, Koh H, Börner K, Storm TA, Davis CR, Kay MA, et al. 2009. Low-level shRNA Cytotoxicity Can Contribute to MYC-induced Hepatocellular Carcinoma in Adult Mice. *Mol Ther* **18**: 161–170.

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

Bennasser Y, Chable-Bessia C, Triboulet R, Gibbings D, Gwizdek C, Dargemont C, Kremer EJ, Voinnet O, Benkirane M. 2011. Competition for XPO5 binding between Dicer mRNA, pre-miRNA and viral RNA regulates human Dicer levels. *Nat Struct Mol Biol* **18**: 323–327.

Berkhout B, Liu YP. 2014. Towards improved shRNA and miRNA reagents as inhibitors of HIV1 replication. *Future Microbiology* **9**: 561–571.

Bernards R, Brummelkamp TR, Beijersbergen RL. 2006. shRNA libraries and their use in cancer genetics. *Nat Meth* **3**: 701–706.

Bhinder B, Djaballah H. 2013. Systematic analysis of RNAi reports identifies dismal commonality at gene-level & reveals an unprecedented enrichment in pooled shRNA screens. *Comb Chem High Throughput Screen* **16**: 665–681.

Bhinder B, Shum D, Djaballah H. 2014a. Comparative analysis of RNAi screening technologies at genome-scale reveals an inherent processing inefficiency of the plasmid-based shRNA hairpin. *Comb Chem High Throughput Screen* **17**: 98–113.

Bhinder B, Shum D, Li M, Ibáñez G, Vlassov AV, Magdaleno S, Djaballah H. 2014b. Discovery of a Dicer-Independent, Cell-Type Dependent Alternate Targeting Sequence Generator: Implications in Gene Silencing & Pooled RNAi Screens. *PLoS ONE* **9**: e100676.

Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin S-F, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, et al. 2007. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biology* **8**: R214.

Boden D, Pusch O, Silbermann R, Lee F, Tucker L, Ramratnam B. 2004. Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins. *Nucl Acids Res* **32**: 1154–1158.

Boele J, Persson H, Shin JW, Ishizu Y, Newie IS, Søkilde R, Hawkins SM, Coarfa C, Ikeda K, Takayama K, et al. 2014. PAPD5-mediated 3′ adenylation and subsequent degradation of miR-21 is disrupted in proliferative disease. *PNAS* **111**: 11467–11472.

Bohnsack MT, Czaplinski K, Gorlich D. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**: 185–191.

Borel F, Kay MA, Mueller C. 2014. Recombinant AAV as a Platform for Translating the Therapeutic Potential of RNA Interference. *Mol Ther* **22**: 692–701.

Boudreau RL, Martins I, Davidson BL. 2008a. Artificial MicroRNAs as siRNA Shuttles: Improved Safety as Compared to shRNAs In vitro and In vivo. *Mol Ther* **17**: 169–175.

Boudreau RL, Monteys AM, Davidson BL. 2008b. Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs. *RNA* **14**: 1834–1844.

Bracht J. 2004. Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA* **10**: 1586–1594.

Brake O ter, Hooft K 't, Liu YP, Centlivre M, Jasmijn von Eije K, Berkhout B. 2008. Lentiviral Vector Design for Multiple shRNA Expression and Durable HIV-1 Inhibition. *Mol Ther* **16**: 557–564.

Brake O ter, Konstantinova P, Ceylan M, Berkhout B. 2006. Silencing of HIV-1 with RNA Interference: A Multiple shRNA Approach. *Mol Ther* **14**: 883–892.

Brennecke J, Stark A, Russell RB, Cohen SM. 2005. Principles of MicroRNA–Target Recognition. *PLoS Biol* **3**: e85.

Brummelkamp TR, Bernards R, Agami R. 2002. Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* **2**: 243–247.

Burroughs AM, Ando Y, de Hoon MJL, Tomaru Y, Nishibu T, Ukekawa R, Funakoshi T, Kurokawa T, Suzuki H, Hayashizaki Y, et al. 2010. A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res* **20**: 1398–1410.

Burroughs AM, Kawano M, Ando Y, Daub CO, Hayashizaki Y. 2011. pre-miRNA profiles obtained through application of locked nucleic acids and deep sequencing reveals complex 5'/3' arm variation including concomitant cleavage and polyuridylation patterns. *Nucl Acids Res* gkr903.

Cai X. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**: 1957–1966.

Castanotto D, Sakurai K, Lingeman R, Li H, Shively L, Aagaard L, Soifer H, Gatignol A, Riggs A, Rossi JJ. 2007. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* **35**: 5154–5164.

Chang K, Elledge SJ, Hannon GJ. 2006. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* **3**: 707–714.

Chang T-C, Yu D, Lee Y-S, Wentzel EA, Arking DE, West KM, Dang CV, Thomas-Tikhonenko A, Mendell JT. 2008. Widespread microRNA repression by Myc contributes to tumorigenesis. *Nature Genetics* **40**: 43–50.

Cheloufi S, Santos CO Dos, Chong MMW, Hannon GJ. 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**: 584–589.

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**: 740–744.

Chung K-H, Hart CC, Bassam S Al-, Avery A, Taylor J, Patel PD, Vojtek AB, Turner DL. 2006. Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucl Acids Res* **34**: e53–e53.

Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, et al. 2010. A Novel miRNA Processing Pathway Independent of Dicer Requires Argonaute2 Catalytic Activity. *Science* **328**: 1694–1698.

Czech B, Zhou R, Erlich Y, Brennecke J, Binari R, Villalta C, Gordon A, Perrimon N, Hannon GJ. 2009. Hierarchical rules for Argonaute loading in Drosophila. *Mol Cell* **36**: 445–456.

Dallas A, Ilves H, Ge Q, Kumar P, Shorenstein J, Kazakov SA, Cuellar TL, McManus MT, Behlke MA, Johnston BH. 2012. Right- and left-loop short shRNAs have distinct and unusual mechanisms of gene silencing. *Nucl Acids Res* gks662.

Davis BN, Hilyard AC, Lagna G, Hata A. 2008. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **454**: 56–61.

Davis BN, Hilyard AC, Nguyen PH, Lagna G, Hata A. 2010. Smad Proteins Bind a Conserved RNA Sequence to Promote MicroRNA Maturation by Drosha. *Molecular Cell* **39**: 373–384.

Denise H, Moschos SA, Sidders B, Burden F, Perkins H, Carter N, Stroud T, Kennedy M, Fancy S-A, Lapthorn C, et al. 2014. Deep Sequencing Insights in Therapeutic shRNA Processing and siRNA Target Cleavage Precision. *Mol Ther Nucleic Acids* **3**: e145.

Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**: 231–235.

Diederichs S, Haber DA. 2007. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**: 1097–1108.

Doench JG, Sharp PA. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev* **18**: 504–511.

Doyle M, Badertscher L, Jaskiewicz L, Güttinger S, Jurado S, Hugenschmidt T, Kutay U, Filipowicz W. 2013. The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. *RNA* **19**: 1238–1252.

Dueck A, Ziegler C, Eichner A, Berezikov E, Meister G. 2012. microRNAs associated with the different human Argonaute proteins. *Nucl Acids Res* **40**: 9850–9862.

Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, Lund E, Dahlberg JE. 2005. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proceedings of the National Academy of Sciences* **102**: 3627–3632.

Elbashir SM, Lendeckel W, Tuschl T. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**: 188–200.

Farazi TA, Juranek SA, Tuschl T. 2008. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **135**: 1201–1214.

Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, Lai DY, Barbosa IAM, Kwon JS, Guan Y, et al. 2013. An Optimized microRNA Backbone for Effective Single-Copy RNAi. *Cell Reports* **5**: 1704–1713.

Fellmann C, Zuber J, McJunkin K, Chang K, Malone CD, Dickins RA, Xu Q, Hengartner MO, Elledge SJ, Hannon GJ, et al. 2011. Functional Identification of Optimized RNAi Triggers Using a Massively Parallel Sensor Assay. *Molecular Cell* **41**: 733–746.

Feng Y, Zhang X, Graves P, Zeng Y. 2012. A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA* **18**: 2083–2092.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**: 806–811.

Flores O, Kennedy EM, Skalsky RL, Cullen BR. 2014. Differential RISC association of endogenous human microRNAs predicts their inhibitory potential. *Nucl Acids Res* **42**: 4629–4639.

Förstemann K, Horwich MD, Wee L, Tomari Y, Zamore PD. 2007. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* **130**: 287–297.

Friedman RC, Farh KK-H, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**: 92–105.

Gagnon KT, Li L, Chu Y, Janowski BA, Corey DR. 2014. RNAi Factors are Present and Active in Human Cell Nuclei. *Cell Rep* **6**: 211–221.

Gantier MP, McCoy CE, Rusinova I, Saulep D, Wang D, Xu D, Irving AT, Behlke MA, Hertzog PJ, Mackay F, et al. 2011. Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucleic Acids Res* **39**: 5692–5703.

Ge Q, Ilves H, Dallas A, Kumar P, Shorenstein J, Kazakov SA, Johnston BH. 2010. Minimal-length short hairpin RNAs: The relationship of structure and RNAi activity. *RNA* **16**: 106–117.

Gregory RI, Yan K, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**: 235–240.

Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. 2006. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**: 537–541.

Grimm D, Wang L, Lee JS, Schürmann N, Gu S, Börner K, Storm TA, Kay MA. 2010. Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *J Clin Invest* **120**: 3106–3119.

Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. 2008. The Vienna RNA Websuite. *Nucleic Acids Res* **36**: W70–W74.

Guil S, Cáceres JF. 2007. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Structural & Molecular Biology* **14**: 591–596.

Guo H, Ingolia NT, Weissman JS, Bartel DP. 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**: 835–840.

Gurtan AM, Lu V, Bhutkar A, Sharp PA. 2012. In vivo structure–function analysis of human Dicer reveals directional processing of precursor miRNAs. *RNA* **18**: 1116–1122.

Gu S, Jin L, Zhang F, Huang Y, Grimm D, Rossi JJ, Kay MA. 2011. Thermodynamic stability of small hairpin RNAs highly influences the loading process of different mammalian Argonautes. *Proc Natl Acad Sci U S A* **108**: 9208–9213.

Gu S, Jin L, Zhang Y, Huang Y, Zhang F, Valdmanis PN, Kay MA. 2012. The Loop Position of shRNAs and Pre-miRNAs Is Critical for the Accuracy of Dicer Processing In Vivo. *Cell* **151**: 900–911.

Gu S, Kay MA. 2010. How do miRNAs mediate translational repression? *Silence* **1**: 11.

Gu S, Zhang Y, Jin L, Huang Y, Zhang F, Bassik MC, Kampmann M, Kay MA. 2014. Weak base pairing in both seed and 3′ regions reduces RNAi off-targets and enhances si/shRNA designs. *Nucl Acids Res* gku854.

Hagan JP, Piskounova E, Gregory RI. 2009. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol* **16**: 1021–1025.

Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**: 293–296.

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. 2001. Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi. *Science* **293**: 1146–1150.

Han J. 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* **18**: 3016–3027.

Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN. 2006. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell* **125**: 887–901.

Hannon GJ, Rossi JJ. 2004. Unlocking the potential of the human genome with RNA interference. *Nature* **431**: 371–378.

Haussecker D. 2015. RNAi Therapeutics: First RNAi Therapeutic Nearing Finish Line. *RNAi Therapeutics*. http://rnaitherapeutics.blogspot.com/2015/04/first-rnai-therapeutic-nearing-finish.html (Accessed July 1, 2015).

Haussecker D, Kay MA. 2015. Drugging RNAi. *Science* **347**: 1069–1070.

He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S, Cordon-Cardo C, Lowe SW, Hannon GJ, et al. 2005. A microRNA polycistron as a potential human oncogene. *Nature* **435**: 828–833.

Heo I, Ha M, Lim J, Yoon M-J, Park J-E, Kwon SC, Chang H, Kim VN. 2012. Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* **151**: 521–532.

Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. 2008. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell* **32**: 276–284.

Heo I, Joo C, Kim Y-K, Ha M, Yoon M-J, Cho J, Yeom K-H, Han J, Kim VN. 2009. TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell* **138**: 696–708.

Herrera-Carrillo E, Harwig A, Berkhout B. 2015. Toward optimization of AgoshRNA moleculesthat use a non-canonical RNAi pathway: Variations in the top and bottom base pairs. *RNA Biology* **12**: 447–456.

Herrera-Carrillo E, Harwig A, Liu YP, Berkhout B. 2014. Probing the shRNA characteristics that hinder Dicer recognition and consequently allow Ago-mediated processing and AgoshRNA activity. *RNA* **20**: 1410–1418.

Hinton TM, Wise TG, Cottee PA, Doran TJ. 2008. Native microRNA loop sequences can improve short hairpin RNA processing for virus gene silencing in animal cells. *J RNAi Gene Silencing* **4**: 295–301.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**: 99–110.

Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore PD. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**: 834–838.

Hutvágner G, Zamore PD. 2002. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**: 2056–2060.

Hwang H-W, Wentzel EA, Mendell JT. 2009. Cell–cell contact globally activates microRNA biogenesis. *Proc Natl Acad Sci U S A* **106**: 7016–7021.

Ibrahim F, Rymarquis LA, Kim E-J, Becker J, Balassa E, Green PJ, Cerutti H. 2010. Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *PNAS* **107**: 3906–3911.

Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* **21**: 635–637.

Janas MM, Wang B, Harris AS, Aguiar M, Shaffer JM, Subrahmanyam YVBK, Behlke MA, Wucherpfennig KW, Gygi SP, Gagnon E, et al. 2012. Alternative RISC assembly: Binding and repression of microRNA–mRNA duplexes by human Ago proteins. *RNA* **18**: 2041–2055.

Jones MR, Quinton LJ, Blahna MT, Neilson JR, Fu S, Ivanov AR, Wolf DA, Mizgerd JP. 2009. Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol* **11**: 1157–1163.

Kaelin WG. 2012. Use and Abuse of RNAi to Study Mammalian Gene Function. *Science* **337**: 421–422.

Katoh T, Sakaguchi Y, Miyauchi K, Suzuki T, Kashiwabara S -i., Baba T, Suzuki T. 2009. Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes & Development* **23**: 433–438.

Kawamata T, Seitz H, Tomari Y. 2009. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol* **16**: 953–960.

Kim K, Lee YS, Carthew RW. 2007. Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. *RNA* **13**: 22–29.

Kim Y-K, Kim VN. 2007. Processing of intronic microRNAs. *The EMBO Journal* **26**: 775–783.

Kiss DL, Andrulis ED. 2010. Genome-wide analysis reveals distinct substrate specificities of Rrp6, Dis3, and core exosome subunits. *RNA* **16**: 781–791.

Knight SW. 2001. A Role for the RNase III Enzyme DCR-1 in RNA Interference and Germ Line Development in Caenorhabditis elegans. *Science* **293**: 2269–2271.

Kok KH, Ng M-HJ, Ching Y-P, Jin D-Y. 2007. Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. *J Biol Chem* **282**: 17649–17657.

Krol J, Loedige I, Filipowicz W. 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics* **11**: 597–610.

Kubowicz P, Zelaszczyk D, Pekala E. 2013. RNAi in Clinical Studies. *Current Medicinal Chemistry* **20**: 1801–1816.

Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genetics* **39**: 673–677.

Landthaler M, Yalcin A, Tuschl T. 2004. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis. *Current Biology* **14**: 2162–2167.

Lavender H, Brady K, Burden F, Delpuech-Adams O, Denise H, Palmer A, Perkins H, Savic B, Scott S, Smith-Burchnell C, et al. 2012. In vitro characterization of the activity of PF-05095808, a novel biological agent for hepatitis C virus therapy. *Antimicrob Agents Chemother* **56**: 1364–1375.

Lee M, Kim B, Kim VN. 2014. Emerging Roles of RNA Modification: m6A and U-Tail. *Cell* **158**: 980–987.

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

Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, Kim VN. 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**: 4051–4060.

Leuschner PJF, Ameres SL, Kueng S, Martinez J. 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep* **7**: 314–320.

Lewis BP, Burge CB, Bartel DP. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**: 15–20.

Li J, Yang Z, Yu B, Liu J, Chen X. 2005. Methylation Protects miRNAs and siRNAs from a 3′-End Uridylation Activity in Arabidopsis. *Current Biology* **15**: 1501–1507.

Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, Kim VN. 2014. Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**: 1365–1376.

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**: 769–773.

Lingel A, Simon B, Izaurralde E, Sattler M. 2004. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**: 576–577.

Liu J. 2004. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* **305**: 1437–1441.

Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R. 2005. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7**: 719–723.

Liu X, Jin D-Y, McManus MT, Mourelatos Z. 2012. Precursor microRNA-programmed silencing complex assembly pathways in mammals. *Mol Cell* **46**: 507–517.

Liu X, Zheng Q, Vrettos N, Maragkakis M, Alexiou P, Gregory BD, Mourelatos Z. 2014. A MicroRNA Precursor Surveillance System in Quality Control of MicroRNA Synthesis. *Molecular Cell* **55**: 868–879.

Liu YP, Haasnoot J, Brake O ter, Berkhout B, Konstantinova P. 2008. Inhibition of HIV-1 by multiple siRNAs expressed from a single microRNA polycistron. *Nucl Acids Res* **36**: 2811–2824.

Liu YP, Karg M, Harwig A, Herrera-Carrillo E, Jongejan A, van Kampen A, Berkhout B. 2015a. Mechanistic insights on the Dicer-independent AGO2-mediated processing of AgoshRNAs. *RNA Biology* **12**: 92–100.

Liu YP, Karg M, Herrera-Carrillo E, Berkhout B. 2015b. Towards Antiviral shRNAs Based on the AgoshRNA Design. *PLoS ONE* **10**: e0128618.

Liu YP, Schopman NCT, Berkhout B. 2013. Dicer-independent processing of short hairpin RNAs. *Nucl Acids Res* gkt036.

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**: 26.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, et al. 2005. MicroRNA expression profiles classify human cancers. *Nature* **435**: 834–838.

Lund E. 2004. Nuclear Export of MicroRNA Precursors. *Science* **303**: 95–98.

Ma H, Wu Y, Choi J-G, Wu H. 2013. Lower and upper stem–single-stranded RNA junctions together determine the Drosha cleavage site. *PNAS* **110**: 20687–20692.

Ma H, Wu Y, Dang Y, Choi J-G, Zhang J, Wu H. 2014a. Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids* **3**: e161.

Ma H, Zhang J, Wu H. 2014b. Designing Ago2-specific siRNA/shRNA to Avoid Competition with Endogenous miRNAs. *Mol Ther Nucleic Acids* **3**: e176.

Maniataki E. 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes & Development* **19**: 2979–2990.

Martinez J, Patkaniowska A, Urlaub H, Lührmann R, Tuschl T. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**: 563–574.

Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD. 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**: 607–620.

McBride JL, Boudreau RL, Harper SQ, Staber PD, Monteys AM, Martins I, Gilmore BL, Burstein H, Peluso RW, Polisky B, et al. 2008. Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: Implications for the therapeutic development of RNAi. *PNAS* **105**: 5868–5873.

McManus MT, Petersen CP, Haines BB, Chen J, Sharp PA. 2002. Gene silencing using micro-RNA designed hairpins. *RNA* **8**: 842–850.

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. 2004. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15**: 185–197.

Melamed Z, Levy A, Ashwal-Fluss R, Lev-Maor G, Mekahel K, Atias N, Gilad S, Sharan R, Levy C, Kadener S, et al. 2013. Alternative Splicing Regulates Biogenesis of miRNAs Located across Exon-Intron Junctions. *Molecular Cell* **50**: 869–881.

Michlewski G, Guil S, Semple CA, Cáceres JF. 2008. Posttranscriptional Regulation of miRNAs Harboring Conserved Terminal Loops. *Molecular Cell* **32**: 383–393.

Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC. 2005. Slicer function of Drosophila Argonautes and its involvement in RISC formation. *Genes Dev* **19**: 2837–2848.

Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, et al. 2006. A Lentiviral RNAi Library for Human and Mouse Genes Applied to an Arrayed Viral High-Content Screen. *Cell* **124**: 1283–1298.

Mohr SE, Perrimon N. 2012. RNAi screening: new approaches, understandings, and organisms. *WIREs RNA* **3**: 145–158.

Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15**: 902–909.

Newman MA, Mani V, Hammond SM. 2011. Deep sequencing of microRNA precursors reveals extensive 3′ end modification. *RNA* **17**: 1795–1803.

Newman MA, Thomson JM, Hammond SM. 2008. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14**: 1539–1549.

Nguyen TA, Jo MH, Choi Y-G, Park J, Kwon SC, Hohng S, Kim VN, Woo J-S. 2015. Functional Anatomy of the Human Microprocessor. *Cell*.

Noland CL, Ma E, Doudna JA. 2011. siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell* **43**: 110–121.

Obernosterer G. 2006. Post-transcriptional regulation of microRNA expression. *RNA* **12**: 1161–1167.

Okamura K, Ishizuka A, Siomi H, Siomi MC. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18**: 1655–1666.

Okamura K, Liu N, Lai EC. 2009. Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. *Mol Cell* **36**: 431–444.

Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* **16**: 948–958.

Park J-E, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature* **475**: 201–205.

Pawlicki JM, Steitz JA. 2008. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **182**: 61–76.

Petri S, Dueck A, Lehmann G, Putz N, Rüdel S, Kremmer E, Meister G. 2011. Increased siRNA duplex stability correlates with reduced off-target and elevated on-target effects. *RNA* **17**: 737–749.

Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grässer FA, van Dyk LF, Ho CK, Shuman S, Chien M, et al. 2005. Identification of microRNAs of the herpesvirus family. *Nat Meth* **2**: 269–276.

Phadke AP, Jay CM, Wang Z, Chen S, Liu S, Haddock C, Kumar P, Pappen BO, Rao DD, Templeton NS, et al. 2011. In vivo safety and antitumor efficacy of bifunctional small hairpin RNAs specific for the human Stathmin 1 oncoprotein. *DNA Cell Biol* **30**: 715–726.

Pillai RS, Bhattacharyya SN, Filipowicz W. 2007. Repression of protein synthesis by miRNAs: how many mechanisms? *Trends in Cell Biology* **17**: 118–126.

Premsrirut PK, Dow LE, Kim SY, Camiolo M, Malone CD, Miething C, Scuoppo C, Zuber J, Dickins RA, Kogan SC, et al. 2011. A Rapid and Scalable System for Studying Gene Function in Mice Using Conditional RNA Interference. *Cell* **145**: 145–158.

PRNewswire. 2015. RNA (miRNA, RNAi & siRNA) Therapy in Oncology Drug Pipeline Update 2015. *PR Newswire*, January 7 http://www.prnewswire.com/news-releases/rna-mirna-rnai--sirna-therapy-in-oncology-drug-pipeline-update-2015-300108056.html (Accessed July 3, 2015).

Rand TA, Petersen S, Du F, Wang X. 2005. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123**: 621–629.

Rao DD, Maples PB, Senzer N, Kumar P, Wang Z, Pappen BO, Yu Y, Haddock C, Jay C, Phadke AP, et al. 2010. Enhanced target gene knockdown by a bifunctional shRNA: a novel approach of RNA interference. *Cancer Gene Ther* **17**: 780–791.

Rao DD, Vorhies JS, Senzer N, Nemunaitis J. 2009. siRNA vs. shRNA: similarities and differences. *Adv Drug Deliv Rev* **61**: 746–759.

Rivas FV, Tolia NH, Song J-J, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L. 2005. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* **12**: 340–349.

Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**: 199–208.

Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**: 58–63.

Shen B, Goodman HM. 2004. Uridine addition after microRNA-directed cleavage. *Science* **306**: 997.

Shimizu S, Kamata M, Kittipongdaja P, Chen KN, Kim S, Pang S, Boyer J, Qin FX-F, An DS, Chen ISY. 2009. Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**: 8.

Silva JM, Li MZ, Chang K, Ge W, Golding MC, Rickles RJ, Siolas D, Hu G, Paddison PJ, Schlabach MR, et al. 2005. Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet* **37**: 1281–1288.

Soifer HS, Sano M, Sakurai K, Chomchan P, Sætrom P, Sherman MA, Collingwood MA, Behlke MA, Rossi JJ. 2008. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucl Acids Res* **36**: 6511–6522.

Song H-W, Bettegowda A, Oliver D, Yan W, Phan MH, de Rooij DG, Corbett MA, Wilkinson MF. 2015. shRNA Off-Target Effects In Vivo: Impaired Endogenous siRNA Expression and Spermatogenic Defects. *PLoS One* **10**. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4366048/ (Accessed July 3, 2015).

Song J-J, Liu J, Tolia NH, Schneiderman J, Smith SK, Martienssen RA, Hannon GJ, Joshua-Tor L. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**: 1026–1032.

Song J-J, Smith SK, Hannon GJ, Joshua-Tor L. 2004. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* **305**: 1434–1437.

Sonia A Melo CM. 2010. A Genetic Defect in Exportin-5 Traps Precursor MicroRNAs in the Nucleus of Cancer Cells. *Cancer cell* **18**: 303–15.

Starega-Roslan J, Krol J, Koscianska E, Kozlowski P, Szlachcic WJ, Sobczak K, Krzyzosiak WJ. 2011. Structural basis of microRNA length variety. *Nucleic Acids Res* **39**: 257–268.

Stegmeier F, Hu G, Rickles RJ, Hannon GJ, Elledge SJ. 2005. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *PNAS* **102**: 13212–13217.

Su H, Trombly MI, Chen J, Wang X. 2009. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev* **23**: 304–317.

Sun C-P, Wu T-H, Chen C-C, Wu P-Y, Shih Y-M, Tsuneyama K, Tao M-H. 2013. Studies of Efficacy and Liver Toxicity Related to Adeno-Associated Virus–Mediated RNA Interference. *Human Gene Therapy* **24**: 739–750.

Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K. 2009. Modulation of microRNA processing by p53. *Nature* **460**: 529–533.

Tan GS, Garchow BG, Liu X, Metzler D, Kiriakidou M. 2011. Clarifying mammalian RISC assembly in vitro. *BMC Mol Biol* **12**: 19.

Tan GS, Garchow BG, Liu X, Yeung J, Morris JP, Cuellar TL, McManus MT, Kiriakidou M. 2009. Expanded RNA-binding activities of mammalian Argonaute 2. *Nucleic Acids Res* **37**: 7533–7545.

Thomson JM. 2006. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development* **20**: 2202–2207.

Thornton JE, Chang H-M, Piskounova E, Gregory RI. 2012. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18**: 1875–1885.

Timmons L, Fire A. 1998. Specific interference by ingested dsRNA. *Nature* **395**: 854–854.

Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, Gherzi R, Rosenfeld MG. 2009. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **459**: 1010–1014.

Ustianenko D, Hrossova D, Potesil D, Chalupnikova K, Hrazdilova K, Pachernik J, Cetkovska K, Uldrijan S, Zdrahal Z, Vanacova S. 2013. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19**: 1632–1638.

Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ. 2008a. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**: 209–213.

Wang Y, Wang YE, Cotticelli MG, Wilson RB. 2008b. A Random shRNA-Encoding Library for Phenotypic Selection and Hit-Optimization. *PLoS ONE* **3**: e3171.

White E, Schlackow M, Kamieniarz-Gdula K, Proudfoot NJ, Gullerova M. 2014. Human nuclear Dicer restricts the deleterious accumulation of endogenous double-stranded RNA. *Nat Struct Mol Biol* **21**: 552–559.

Winter J, Diederichs S. 2011. Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization. *RNA Biology* **8**: 1149–1157.

Wu H, Ma H, Ye C, Ramirez D, Chen S, Montoya J, Shankar P, Wang XA, Manjunath N. 2011. Improved siRNA/shRNA functionality by mismatched duplex. *PLoS ONE* **6**: e28580.

Yang J-S, Maurin T, Lai EC. 2012. Functional parameters of Dicer-independent microRNA biogenesis. *RNA* **18**: 945–957.

Yang J-S, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC. 2010. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *PNAS* **107**: 15163–15168.

Yang J-S, Smibert P, Westholm JO, Jee D, Maurin T, Lai EC. 2014. Intertwined pathways for Argonaute-mediated microRNA biogenesis in Drosophila. *Nucl Acids Res* **42**: 1987–2002.

Yang W, Chendrimada TP, Wang Q, Higuchi M, Seeburg PH, Shiekhattar R, Nishikura K. 2006. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature Structural & Molecular Biology* **13**: 13–21.

Yi R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development* **17**: 3011–3016.

Yoda M, Cifuentes D, Izumi N, Sakaguchi Y, Suzuki T, Giraldez AJ, Tomari Y. 2013. Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Rep* **5**: 715–726.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**: 25–33.

Zeng M, Kuzirian MS, Harper L, Paradis S, Nakayama T, Lau NC. 2013. Organic small hairpin RNAs (OshR): A do-it-yourself platform for transgene-based gene silencing. *Methods* **63**: 101–109.

Zeng Y, Cullen BR. 2003. Sequence requirements for micro RNA processing and function in human cells. *RNA* **9**: 112–123.

Zeng Y, Cullen BR. 2004. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucl Acids Res* **32**: 4776–4785.

Zeng Y, Wagner EJ, Cullen BR. 2002. Both Natural and Designed Micro RNAs Can Inhibit the Expression of Cognate mRNAs When Expressed in Human Cells. *Molecular Cell* **9**: 1327–1333.

Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**: 57–68.

Zhang X, Zeng Y. 2010. The terminal loop region controls microRNA processing by Drosha and Dicer. *Nucl Acids Res* **38**: 7689–7697.