Using Engineered microRNAs as Vectors for Animal RNA Interference: Promises and Challenges

Jerry S. Chen¹, Robert W. Zeller²

¹Computational Science Research Center, San Diego State University, San Diego, USA
²Department of Biology, San Diego State University, San Diego, USA
Email: rzeller@mail.sdsu.edu

Received 17 December 2013; revised 4 February 2014; accepted 21 February 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract
microRNAs are post-transcriptional regulators of gene expression that recruit RNA silencing complexes to target transcripts to prevent translation and promote their degradation. Experimental studies suggest that microRNA binding to target transcripts can result in as much as a 90% decrease in gene expression. Because of this feature, the microRNA pathway has been utilized as a vehicle for potent RNA interference (RNAi). In recent years, significant advances have been made in engineering artificial microRNA vectors for RNAi in a number of biological systems, with the most progress in plants but also some success in mouse and human cell lines. In this mini-review, we provide a brief discussion of the potential of this technology in comparison with other RNAi strategies, and the current challenges in the design of microRNA-based RNAi vectors, particularly for animal systems.

Keywords
microRNAs; RNA Interference; RNAi

1. Introduction
microRNAs (miRNAs) are small RNA molecules that play important developmental roles in both animals and plants, known clinically for their proposed role in preventing cancer formation [1]-[3]. They were first identified in 1993 by the laboratories of Victor Ambros and Gary Ruvkun [4] [5]. Using the nematode worm C. elegans, Ambros and Ruvkun showed that the gene product of a regulatory molecule called lin-4 was necessary for suppressing the activity of a transcription factor (TF) protein called LIN-14 during early larval development. Inte-
restingly, the product of lin-4 was not a protein as would typically be expected for a molecule with a repressive function, but rather was a tiny RNA molecule less than 70 nucleotides (nt) long. It was then discovered that the sequence of the lin-4 RNA product was highly complementary to a region within the 3’ untranslated region (3’UTR) of the LIN-14 messenger RNA (mRNA) transcript. They then suggested that perhaps this RNA-RNA interaction controls the expression of the LIN-14 protein. A few other studies soon followed [6]-[9] but it was not until the complete sequencing of several genomes (including the human genome) at the turn of the century [10]-[13] that microRNA biology blossomed. A flurry of research between 2000 and 2003 demonstrated that these microRNAs were not just chance occurrences, but were in fact abundant in the genome and pervasive throughout the plant and animal kingdom [14]-[36]. This sparked an international scientific endeavor to identify and discover the function of these curious runts of the genome, which continues to this day [3] [37].

2. Using microRNAs as Vectors for RNA Interference

2.1. How microRNAs Are Processed

microRNAs, like protein-coding genes, are encoded in the genome and transcribed by RNA polymerase II. They can reside as stand-alone genes in independent genetic loci or reside within the introns of host genes. In rare cases, they may even bury themselves within the exons of protein-coding genes [38] [39]. Their processed hairpin transcripts are small, only ~70 base pairs (bp) in length, explaining why they evaded detection in the early days of genome research.

What facilitated the widespread discovery of microRNAs was the resolution of their RNA structure. After the primary miRNA is transcribed from the genome, it folds into a stem-loop structure, which is trimmed into a distinguishable ~70 bp hairpin-shaped pre-miRNA molecule by the RNase III enzyme Drosha [38] [39] (Figure 1). The stems of the pre-miRNA are highly complementary, enabling scientists to find predicted microRNAs by searching for genomic regions with adjacent inverse complementary sequences separated by a small gap that corresponds to the pre-miRNA loop. Since regions of the genome may have consecutive inverse complementary regions by chance, these computational studies generally predict more microRNAs than are found experimentally [18] [40]. Nonetheless, these initial computational studies guided subsequent experimental efforts, which used small-scale methods such as in situ hybridization [41]-[43] and Northern blot [16] [19] [35] as well as large-scale methods such as RNA deep sequencing [32] [34] [44] [45] and miRNA microarrays [35] [46]. To date, the current numbers of identified microRNAs range from 200 in C. elegans [47] [48] to over 1000 in humans [48] [49]. Interestingly, whereas the number of protein-coding genes increases at only a modest rate, the number of microRNAs in the genome increases quite significantly in proportion with the complexity of the organism, indicating that the increased complexity of higher organisms may in large part be the result of microRNA expansion [50].

After trimming by Drosha, the 70 nucleotide hairpin pre-miRNA is then shuttled from the nucleus out into the cytoplasm by the nuclear transport receptor Exportin 5, where the loop of the hairpin is then cut by the RNase III enzyme Dicer [38] [39]. From the resultant ~20 - 24 bp RNA duplex, one strand becomes the mature miRNA product. The other strand, known as the miRNA*, is generally thought to be degraded, although recent reports suggest that in a few cases they may also have regulatory capability [51] [52]. The mature miRNA is loaded into a RNA silencing complex containing the RNA-binding protein Argonaute and serves as a guide strand to target and bind partially complementary sites in 3’UTR regions of target mRNAs to regulate their expression [3] [38] [39]. Evidence also exists that Argonaute-complexes frequently bind within coding regions, suggesting that miRNAs may also target these regions as well [44] [45].

2.2. Engineered microRNA Vectors for RNAi: The Promises

RNA interference was first achieved by introduction of synthetic long double-stranded RNA into cells or animals, which transiently produce siRNA strands complementary to the desired target gene transcript [53] [54]. RNA interference was first demonstrated in C. elegans [53] and soon after in mammalian cell lines [54] [55]. In more recent years, RNAi has also been used as a therapeutic tool to knock-down the expression of genes implicated in the pathophysiology of many cancers, viral infections, and other diseases [56]-[61].

For more stable and controlled expression of siRNAs in model organisms, vectors have been developed for transcribing short-hairpin RNAs (shRNAs) under the control of RNA polymerase III promoters such as U6 or...
Figure 1. microRNA hairpin structure. Shown here is the pre-miRNA hairpin structure for *Ciona* miR-155, with the identified mature miR (red) and miRstar (blue) indicated. pre-miRNA hairpins can vary in length from 60 to ~100 nt, and sometimes have tail flanking regions that are also complementary [36]. The ends of the miR-miR* duplex are cleaved by the RNAse III enzyme Dicer, which leaves overhangs on both ends.

**H1** [62] [63], which can either be expressed via extra-chromosomal arrays or directly integrated into the genome using retroviral transduction [62] [64]. Many labs have utilized this technology to create shRNA libraries and stable transgenic lines [62] [65]. However, this strategy does not work in all organisms and in some systems suffers from low efficiency and high off-target effects [66]. Furthermore, since it appears that stable shRNA expression is only achievable using RNA polymerase III promoters, which are constitutively expressed at a certain level in all cell types, the ability to control the timing and level of shRNA expression is currently not possible [62].

An alternative approach is to use microRNA hairpin transcripts as a vector for generating silencing RNAs [62] [67]-[69] (Figure 2). Using this technique, the microRNA hairpin undergoes endogenous processing by the microRNA biogenesis machinery, producing an siRNA-like strand which then targets a desired gene transcript. The hairpin is artificially designed such that the mature and miRNA* strands are replaced by a double-stranded miRNA/siRNA-like duplex, and the entire hairpin cassette is cloned into an appropriate RNA-polymerase II promoter-driven vector (Figure 2). In order to more accurately mimic miRNA duplexes and distinguish themselves from double stranded siRNAs, a bulge is sometimes introduced in the strand replacing the miRNA* so that the opposing strands of the duplex are not perfectly complementary [54]. The result is a vehicle for RNAi that uses the miRNA pathway for processing and can produce tissue-specific miRNA-like small RNAs. Using
transcripts containing miRNA clusters, one can engineer a multi-cassette vector producing multiple miRNA-like small RNAs to target multiple transcripts. Furthermore, since this strategy essentially mimics the processing of endogenous miRNA transcripts, introduction of a miRNA-based vector into cells may circumvent unwanted biological side effects such as type 1 interferon responses that sometimes occur when introducing double-stranded siRNA [68] [70]. RNAi vectors using engineered microRNAs have been successfully developed in a number of systems, most notably in plants [71]-[74], algae [75] [76] and mouse and human cell lines [55] [62]. Collectively, these studies have reported high-specificity of targeting and better knockdown efficiencies than shRNAs. In agreement with previous miRNA studies, miRNA-based RNAi vectors used in human cell lines were shown to result in as much as ~80% knockdown of gene expression, and the degree of RNAi knockdown is dose-dependent [62] [68]. Notably, it has been shown that efficient RNAi knockdown can result in observable loss-of-function phenotypes [64]. Importantly though, because miRNA target binding in plants requires near-perfect complementarity while in animals requires only partial complementarity, the potential for off-target effects is an important factor to consider when designing and using artificial miRNA RNAi vectors in animal systems. We discuss these challenges below.

2.3. Engineered microRNA Vectors for RNAi: The Challenges

One of the primary challenges in using microRNAs for RNA interference in animals is in determining the precise rules for how microRNAs bind to their targets. In plants, microRNAs bind to target sites with near-perfect or perfect complementarity, and therefore design of corresponding RNAi vectors is more straightforward. However, in animals, microRNAs can bind to targets with only partial complementarity, and we still do not know exactly how a microRNA binds to an mRNA target. Through experimental studies over the past decade, it is becoming increasingly clear, however, that there are three large classes of targets based on mRNA binding within the first eight nucleotides of the miRNA, which we refer to as the “seed” region [3]. These are called 8 mer, 7 mer-m8 and 7 mer-A1 seed targets, respectively (Figure 3(A)) [3] [40] [43] [77] [78]. If we merely searched for these kinds of sites within 3' UTR of mRNA transcripts, we would predict that a single microRNA would target on average about 2000 transcripts in human, around 800 per miRNA in simple chordates, and
about 300 - 400 in worms and fruit flies [42] [43] [78]-[80]. Many computational target prediction programs are based on finding seed sites [43] [77] [79]. However, it is generally thought that these are overestimates, and that there are many other biological factors involved such as 3’UTR structure and local thermodynamics which may nullify many of these predicted targets [3] [81]. Although it seems very likely that any mRNA containing a seed site in its 3’UTR is a bona fide target, the lack of precise large-scale protein-based methods hinders our ability to solve this problem definitively. However, from many target expression studies, we do know that the level of downregulation among verified seed targets can vary widely, ranging from as little as 10% to over 90% downregulation [43] [77]. To complicate matters, there are several other types of target binding that have been reported to be functional in some, but not all cases. These include mRNA binding with the presence of a single G:U wobble pair within the seed, binding of only miRNA nucleotides 2 - 7, binding of the center of the miRNA, and compensatory binding where complementarity of only 4 - 5 base pairs of the miRNA seed is compensated by extensive binding of the rest of the miRNA [3] [4] [40] [82]-[84] (Figure 3(B)). More recently, an alternative approach to finding miRNA targets utilizes a biochemical assay to isolate and sequence Argonaute-bound messenger transcripts, which has expanded our understanding of base-pairing rules and interestingly has shown that a large percentage of targets are bound in exon regions [44] [45] [84] [85]. In particular, a ligation-based method has been developed for directly associating miRNAs with their RNA targets [84]. Although the efficiency of this protocol is currently very low (only 2% of sequenced reads are ligated miRNA-RNA hybrid reads), this method holds exciting promise for uncovering miRNA target binding rules and improving microRNA-based RNAi technology.

A second major challenge is in considering the regulation of microRNA biogenesis. Once miRNAs are transcribed, the hairpin forms within a larger primary transcript. It has been demonstrated that sequences within the flanking 5’ and 3’ tails are required for proper Drosha processing [86]. Sequences in the 3’ tail have been shown to contain miRNA binding sites, suggesting that the 3’ tails may serve as a regulatory region akin to 3’UTRs for messenger RNAs [87] [88]. Furthermore, both the tail and loop regions contain binding sites for regulatory RNA binding proteins [89] [90]. Therefore, when expressing miRNA-based RNAi vectors outside of their endogenous context, potential regulation in these loop and tail regions must be considered. Even for the miRNA duplex, one must consider the potential production of miRNA isoforms (isomiRs) that differ in either the 5’ or 3’ end [91] [92], which may be a source of unexpected off-target effects. Since many facets of the microRNA biogenesis pathway are still poorly understood [38], there is currently still an element of trial-and-error in designing proper miRNA RNAi vectors.

3. Conclusion

Despite intensive research over the past decade, the field of microRNA biology is still in its infancy and many
fundamental questions remain unanswered. There are still many gaps in our understanding of the molecular machinery involved in microRNA biogenesis, processing, targeting, and degradation. Nonetheless, microRNAs are naturally produced in almost all model systems, have the demonstrated ability to induce near-complete knockdown of target transcript and/or protein levels, and offer the versatility necessary to be used as vehicles for RNA interference in animals.

Acknowledgements

J.S.C. thanks V. Hurless for helpful discussions. This work was supported by National Science Foundation Grants IOS-0347937 and IOS-0951347 to R.W.Z.

References


RNA-Binding Protein and microRNA Target Sites by PAR-CLIP. *Cell, 141*, 129-141. [http://dx.doi.org/10.1016/j.cell.2010.03.009](http://dx.doi.org/10.1016/j.cell.2010.03.009)


