An overview of microRNAs

Scott M. Hammond
Department of Cell Biology and Physiology Lineberger Comprehensive Cancer Center University of North Carolina Chapel Hill, NC 27599 919-843-2366 hammond@med.unc.edu

Abstract

The discovery of the first microRNA (miRNA) over 20 years ago has ushered in a new era in molecular biology. There are now over 2000 miRNAs that have been discovered in humans and it is believed that they collectively regulate one third of the genes in the genome. miRNAs have been linked to many human diseases and are being pursued as clinical diagnostics and as therapeutic targets. This review presents an overview of the miRNA pathway, including biogenesis routes, biological roles, and clinical approaches.

Graphical Abstract

Keywords

miRNA; microRNA; Dicer; Drosha; Argonaute; RNAi

1. miRNAs are small noncoding regulatory RNAs

The first miRNA was discovered over 30 years ago in the nematode Caenorhabditis elegans with the identification of the developmental regulator lin-4[1]. Originally believed to be a conventional protein coding gene, the Ruvkun and Ambros labs made the startling discovery...
that lin-4 did not code for a protein but instead coded for a 22 nucleotide regulatory RNA[2, 3]. They demonstrated that the lin-4 RNA could base pair with the mRNA for another gene in the C. elegans developmental network, lin-14, and control the production of the LIN-14 protein[3]. This discovery would have had limited significance outside of the C. elegans research community had a second miRNA, let-7, not been discovered[4]. More importantly, let-7 is conserved across many organisms, including humans, suggesting that this class of small regulatory RNAs has a more general role in biology[5]. The other stunning development, occurring around the same time, was the discovery of the RNAi pathway; specifically, the ~21 nucleotide RNA triggers of the silencing machinery[6-9]. These two pathways have since been shown to be different arms of the same gene silencing pathway[10-12]. Subsequently, many thousands of miRNAs have been discovered in many organisms, and there are currently 2588 annotated miRNAs in the human genome[13]. Since each miRNA is predicted to regulate the expression of hundreds of target genes, the miRNA pathway as a whole is a critical mechanism for gene expression control[14].

2. The canonical miRNA biogenesis pathway

All miRNA families undergo a series of biogenesis steps that convert the primary miRNA transcript into the active, ~22 nucleotide mature miRNA (see Figure 1). The mature miRNA is loaded into the RNA induced silencing complex (RISC) where it directs the complex to target mRNAs, leading to translational repression and target mRNA degradation. This section will cover the canonical miRNA biogenesis pathway that is responsible for the maturation of most miRNA families. For simplicity we will focus on the mammalian enzyme components. Several exceptions to this canonical pathway have been described for unique miRNA families, and are discussed in a later section.

2.1. Transcription

miRNA genes are located throughout the genome[15]. Many miRNA genes are noncoding genes whose sole transcriptional product is the miRNA. In other cases the miRNA is located within an intron or untranslated region (UTR) of a protein coding gene. Figure 2 exemplifies a cluster of three miRNAs located within an intron of the protein coding gene MCM7[16, 17]. The defining feature of all miRNA genes is the stem-loop precursor RNA structure, with one (or sometimes both) strands of the stem the source of the mature miRNA[18-20]. In the case of MCM7 there are three clustered miRNA stem loops, each leading to a distinct mature miRNA with a unique target set[21]. miRNAs are not generally located within coding exons, as excision of the miRNA would lead to loss of the protein coding transcript.

Transcription of the miRNA host gene by RNA polymerase II leads to the pri-miRNA, or primary miRNA transcript[22, 23]. The pri-miRNA is typically spliced, capped, and polyadenylated in a similar manner to protein coding mRNAs[24]. While most miRNA promoters have not been characterized in detail, the few examples studied are similar in structure to protein coding gene promoters, including control elements and histone marks[25-33].
2.2. Processing by Drosha and Dicer

The pri-miRNA requires two endonuclease processing steps before it becomes a mature, active miRNA[34]. The first processing step occurs during or subsequent to transcription of the pri-miRNA by the enzyme Drosha[35-38]. The RNA binding protein DGC8 is associated with Drosha and is required for cleavage of the pri-miRNA[39-41]. Drosha cleavage releases the stem-loop precursor from flanking pri-miRNA transcript sequences. The precursor is exported out of the nucleus in a Ran-GTPase dependent manner by Exportin5[42-44]. In the cytoplasm the second processing step occurs. The endonuclease Dicer cleaves the loop region of the precursor releasing the mature miRNA[10-12, 45]. Like Drosha, Dicer is also associated with an RNA binding protein, in this case TRBP[46-50]. The product of the Dicer reaction is a duplex RNA of approximately 21 nucleotides length (depending on the miRNA). One strand of the duplex is loaded into RISC as a mature miRNA. The other strand, termed the star strand, is typically degraded. An asterisk (star) is appended to the miRNA name to designate this star strand, e.g. miR-145*. However, for some miRNAs both strands may be loaded into RISC at similar frequencies. In this case, the strand from the 5’ end of the stem-loop is termed “5p” and the 3’ strand the “3p”. In fact, while RISC loading may strongly favor incorporation of one strand, recent next generation sequencing (NGS) efforts have demonstrated a small fraction of star strand loaded for essentially all miRNA families[51]. Further complicating the nomenclature, some miRNAs show different strand usage depending on cell type or biological state[52]. For these reasons, 5p/3p naming schemes are being increasingly used rather than the arbitrary mature/star nomenclature.

The specific nucleotide locations of Drosha and Dicer cleavage are usually narrowly defined, leading to mature miRNAs with clearly defined terminal ends. Some miRNAs, however, have heterogenous cleavage sites, leading to multiple “isomiRs” of the mature miRNA[53, 54]. This heterogeneity could lead to differential target repression and distinct biological activities. IsomiR preference can vary by cell type and isomiR switching has been detected in disease[52, 55, 56]. The molecular mechanism for isomiR switching is unclear.

2.3. RISC loading and target repression

The ultimate fate of the miRNA is to be incorporated into RISC (or miRISC). The exact composition of this protein complex is not clear but contains the essential protein Argonaute, of which four family members have been identified in humans (Ago1-4)[57, 58]. After Dicer cleavage, the miRNA duplex is loaded into RISC and the star strand is removed by one of several possible mechanisms. If the miRNA duplex has complementarity in the central region, the star strand can be cut by Ago2 and further degraded by the nuclease complex C3PO[59-62]. This is the mechanism for RISC loading for the related siRNA pathway. Most miRNA duplexes, however, lack central complementarity and therefore cannot participate in star strand cleavage. These miRNA duplexes rely on strand unwinding, and several helicases have been described to possess this activity[63-67].

Argonaute directly binds the mature miRNA and seeks target mRNAs that have complementarity to the miRNA. In particular, nucleotides 2-7 of the miRNA, termed the “seed” region, are important for target association[68-70]. The 3’ end of the miRNA can also
contribute to target recognition, and centrally matched targets have been identified[71-74]. If complementarity exists in the central region of the miRNA (nucleotides 9-11) then the mRNA target can be cleaved via the endonuclease activity of Ago2[75, 76]. Most miRNA targets in humans, however, lack this central sequence match and are not directly cleaved by Ago2. Rather, Argonaute is recruited to a complex containing GW182 (TNRC6A/B/C) within cytoplasmic P bodies where translational repression occurs. The CCR4-NOT deadenylase complex is recruited to RISC and this facilitates removal of the poly(A) tail and eventual degradation of the mRNA target [77-79].

2.4. Alternative pathways

While the majority of identified miRNAs are produced by the canonical pathway there are several examples of alternative biogenesis. The most common variation is the mirtron class of miRNAs[80-83]. These miRNAs, located in introns at the exon junction site, bypass the Drosha step. Instead, the precursor ends are formed by the splicing cleavage events. The resultant precursor stem loop is processed by Dicer and loaded into RISC. Dicer independent miRNA biogenesis has also been described[84-86]. The best studied example is miR-451. After transcription and Drosha cleavage the precursor directly binds Ago2, instead of Dicer, and Ago2 cleaves the star strand (like it would cleave a target RNA). The remaining loop sequence is trimmed back to yield the pre-loaded, mature miR-451. Other miRNAs may follow this same pathway[87]. These non-canonical pathways are of particular relevance when interpreting data from Dicer or Drosha/DGCR8 knockout studies.

While the aforementioned RNAs are mature miRNAs that are generated through non-canonical routes, there are other classes of small RNAs that are not miRNAs, yet are produced by Drosha and/or Dicer activity and bound to RISC[88]. The best characterized class of these RNAs is the tRNA derived small RNA fragments (tRF or tsRNAs). These RNAs are a similar size to miRNAs and can be moderately abundant. Some tsRNAs are produced by Dicer cleavage of mature tRNAs and are loaded into RISC[89, 90]. Another class is derived from outside the mature tRNA and is generated by the tRNA splicing reaction[89, 90]. Examples of these RNAs are able to repress gene expression similar to conventional miRNAs, and have been shown to have essential cellular functions[90].

Further complicating the model, several studies have suggested that miRNAs can function independent of the RISC complex. A quantitative analysis of the total miRNA population in a cell, compared to the RISC protein copy number, suggested that a large fraction of miRNAs are not associated with RISC[91]. The functional roles of these RISC-free miRNAs are not known. However, one miRNA, miR-328, has been reported to be in a complex with the RNA binding protein hnRNP E2[92]. This miRNA functionally blocks association of the RNP with its target mRNA C/EBPα, enhancing translation of the C/EBP protein. It should be noted that these non-canonical pathways appear to be rare events. For example, miR-328 predominantly functions as a canonical miRNA within the RISC complex, directing translational repression of target mRNAs[93-97].
3. miRNA target identification

Defining target sets for miRNAs is important for several reasons. For biologists, defining the target set of a miRNA is key to understanding its biological role. For scientists developing miRNA therapeutics, validated targets provide the best biomarker(s) for determination of the efficacy of a miRNA mimic or inhibitor. The identification of miRNA targets has followed three general approaches: bioinformatic target prediction, biochemical isolation of miRNA/mRNA complexes, and transcriptomic/proteomic analysis. The three approaches will be briefly summarized.

3.1. Bioinformatic target prediction

Since miRNAs base pair with target mRNAs using standard Watson-Crick rules, this fact should make bioinformatic target identification straightforward. Unfortunately, the most important determinant of target binding is the seed sequence of the miRNA which is only 6 nucleotides in length. This will lead to a great number of candidate targets, many of which are false negatives. Therefore, all bioinformatic target prediction algorithms use additional factors to improve accuracy\[98\]. Since the best characterized targets are in mRNA 3′ untranslated regions, many algorithms limit target sites to this region. Other factors are employed, including sequence conservation, target site accessibility, flanking sequence determinants, and compensatory pairing outside the seed region. Current approaches have also used machine learning algorithms that incorporate validated targets in the learning sets[98]. A number of algorithms have been developed, with TargetScan, miRanda, and PicTar perhaps the most popular[69, 99-101]. In general, bioinformatic approaches are a good starting point for miRNA analysis and many research labs make use of them[102, 103].

3.2. Biochemical target identification

Several related approaches have used physical association of miRNA/RISC complexes with target mRNAs to isolate and identify targets. These are based on immunoprecipitation of RISC using anti-Argonaute antibodies, with or without prior RNA crosslinking, and definition of bound target RNAs by microarray or NGS profiling[104-107]. Crosslinking before cell lysis is generally preferred since artifactual RNA associations have been observed during cell lysis[108]. An alternative approach is isolation of specific biotinylated miRNAs followed by target identification[109, 110]. This has the advantage of capturing targets of a single known miRNA, but requires ectopic introduction of the biotinylated miRNA. While these physical approaches have been successfully used to define binding partners of miRNA complexes, not all bound mRNA targets may be biologically repressed targets. Studies have reported that some target sites, often in coding sequences of mRNAs, are bound to Argonaute but are not repressed targets[104]. Like all approaches, validation of individual targets is essential.

3.3. Omics-based strategies for target identification

The third general approach to target identification is a proteomic or transcriptomic analysis of cells/tissues in the presence and absence of a miRNA. Quantitative proteomic analysis directly measures the effect of a miRNA on protein production, and is possibly more
reflective of the true target set, but is also technically challenging. Since most miRNA targets have reduced mRNA steady state levels, the simpler transcriptome analysis can be performed[111]. This can be done by microarray profiling and several analysis tools are available[112-115]. An example of this approach is the identification of targets of the neutrophil-specific miRNA miR-223[116]. Neutrophils were isolated from wild type and miR-223 knockout mice. Microarray and quantitative mass spectrometry were performed, and differences in mRNA and protein content was used to define targets of miR-223. It should be noted that the candidate target sets will also contain downstream secondary targets, requiring validation of individual targets.

4. Regulation of the miRNA pathway

Early work in the field focused on transcriptional regulation of miRNA host genes[25, 26, 117, 118]. For example, the cardiac-specific miRNA families miR-1 and miR-133 are transcriptionally regulated by well established cardiomyocyte transcription factor networks via SRF, MEF2, and MyoD[119-123]. Accordingly, enhancer elements for these factors have been identified in upstream regions of the miR-1/133 host genes. Similarly, miRNAs that are located in introns of protein coding genes are often regulated by the same transcription factor networks that control the protein coding mRNA. In general, approaches that have been used for mRNA promoter analysis will work similarly for miRNA promoters.

In addition to transcriptional regulation, miRNA production is often controlled during post-transcriptional biogenesis steps[34]. A well studied example of this is the regulation of the let-7 family during differentiation of embryonic cells. In stem cells the levels of mature let-7 are virtually undetectable. During differentiation, let-7 expression increases over a hundred fold, becoming one of the most abundant miRNA families in differentiated cells of a variety of cell types. Interestingly, transcription of the let-7 host genes does not increase significantly[124]. Rather, let-7 maturation is blocked in embryonic cells by the RNA binding protein Lin28[125-127]. This protein directly binds to the loop region of the pri-miRNA and precursor and inhibits Drosha and Dicer cleavage, respectively[128-130]. In a secondary repressive mechanism, Lin28 directs the uridyl transferase enzyme Tut4 to add an oligouridine tail to let-7 precursors, marking them for degradation[131-134]. During cellular differentiation, Lin28 expression is shut down allowing maturation of abundant let-7 pri-miRNAs. More recently, other RNA binding proteins have been linked to the regulation of miRNA biogenesis, with activating and repressing roles described[135-138]. Regulated degradation of specific miRNAs has also been described[139].

5. Biological roles of miRNAs

The biological importance of miRNAs in mammalian development was first demonstrated with the generation of mouse models deficient for Dicer and DGCR8 (Drosha is involved in other RNA metabolic pathways while DGCR8 is specific for miRNA biogenesis, thus the DGCR8 knockout is more informative). Loss of either step in miRNA biogenesis results in embryonic lethality[140, 141]. Similarly, tissue-specific knockout of either gene leads to developmental defects in that tissue, underscoring the requirement for miRNA function in most tissue types[142]. Of course, knockout of these genes leads to loss of all miRNAs.
(except for the rare non-canonical miRNAs, described above); therefore, it is difficult to assign roles for specific miRNAs in the development of those tissues. This has been addressed by miRNA knockout mouse models[142, 143]. In general, individual miRNAs are not required for specification of individual tissues. For example, mice lacking the cardiac specific miRNA miR-208 still develop a heart[144, 145]. Rather, miRNAs are often required for maintaining tissue homeostasis. In the heart example, miR-208 deficient animals have defects in stress response and exhibit cardiac hypertrophy. miRNAs may be necessary for maintaining tissue differentiation state, since many tissue-restricted miRNAs are reduced in diseases, including cancer (see below).

Transgenic overexpression studies are another approach for the identification of the biological roles of individual miRNAs. These studies have the disadvantage of overexpression artifacts that could lead to misinterpretation of data. However, they have been used to uncover biological roles that have not been detected in knockout studies due to miRNA family redundancy. Furthermore, miRNA expression is often altered in disease, and overexpression studies can mimic this pathological condition (see below). The above-mentioned approaches, along with miRNA target network identification, have revealed roles for the miRNA pathway in the development and/or function of most tissues in the body.

miRNAs have been linked to numerous aspects of neuronal function, including developmental remodeling, amygdala-dependent memory formation, dendritic spine development, and post-mitotic neuron cell survival [146-149]. miRNAs have been linked to the function of many hematopoietic lineages including neutrophils, lymphocytes, erythrocytes, megakaryocytes, an hematopoietic stem cells [17, 150-161]. These examples are a limited subset of the described functions of miRNAs in normal human biology.

6. Dysregulation of miRNAs in disease

Gene expression profiling studies have demonstrated alterations in miRNA expression in a wide range of human disease. In many cases, functional studies have linked miRNA dysregulation as a causal factor in disease progression. This section will summarize some of the disease states that have been linked to miRNA alterations.

6.1. Cancer

Alterations in miRNA expression have now been demonstrated in many cancer types. Early studies were performed using microarray, RT-PCR, and bead-based hybridization (Luminex) platforms, while more recent studies have used NGS-based profiling[117, 162-167]. miRNAs have been identified that are elevated in cancer, for example miR-21 and the miR-17-92 cluster, while other miRNA families are frequently downregulated in cancer, including the let-7 and miR-34 families[168]. Tissue specific miRNAs are often reduced in cancer: for example, the liver-specific miR-122 in hepatocellular carcinoma[169, 170]. This supports the hypothesis that many miRNA families reinforce cellular differentiation state, and reductions in these miRNAs may allow pathological reduction in differentiation or cell state. Functional studies have supported the role of some miRNAs in promoting or preventing cancer development and progression. A well established example is the miR-17-92 cluster. This cluster of six miRNAs is elevated in many cancer types, most notably in several lymphomas and leukemias[171]. Many studies have demonstrated that
this miRNA cluster can promote cancer when ectopically expressed, often in cooperation with the oncogene c-Myc[172-175]. miR-34 is another notable cancer-associated miRNA family[176]. Comprised of three family members, these miRNAs are directly transactivated by p53 and function in parallel with p21 and Bcl2 family proteins to promote cell cycle arrest and apoptosis. Interestingly, this miRNA is at the focus of a frequent deletion in neuroblastoma. The functional importance of these miRNAs in cancer pathogenesis presents a strong case for their development as therapeutic targets[177].

Many other miRNA families are altered in cancers, and large scale profiling studies have uncovered signatures for cancer types and subtypes[178]. There is significant interest in developing miRNA signatures as clinical diagnostics.

6.2. Metabolic disease

miRNAs have been linked to several metabolic pathways, including cholesterol and fatty acid metabolism and transport, pancreatic islet function, and glucose metabolism[179]. miR-122 was introduced in an earlier section as a liver-specific miRNA, and is in fact the most abundantly expressed miRNA in hepatocytes. Targeted deletion of miR-122 in mouse models leads to reduced serum cholesterol and triglyceride levels[180, 181]. This effect was also observed in mouse and primate studies whereby miR-122 was inhibited by an antisense oligonucleotide[182, 183]. The effect on serum cholesterol is due in part to reduction in cholesterol biosynthetic enzyme expression in hepatocytes. miR-33 has been linked to serum high density lipoprotein (HDL) and triglyceride levels, though it is not clear if this is a hepatocyte-specific event or if expression in macrophages, among other cells, is controlling lipid metabolic regulatory networks[184-187]. Both of these miRNAs are being investigated as potential therapeutic targets.

Several miRNA families have been linked to glucose metabolism. miR-375 is highly expressed in pancreatic islet cells and regulates a large number of genes involved in islet cell proliferation and function[188, 189]. Let-7 has also been linked to glucose metabolism. Transgenic overexpression of let-7 in the pancreas leads to reduced glucose tolerance, while overexpression of the let-7 inhibitory protein Lin28 improves glucose uptake[190, 191]. The miR-103/107 family has also been functionally linked to glucose metabolism[192].

6.3. Viral pathogenesis

Several interesting connections have been made between miRNAs and virus biology. To date, 522 viral-encoded miRNAs have been discovered, most notably in the herpesvirus family[13]. While these miRNAs are not essential for virus replication, at least in the examples studied so far, there is evidence that miRNAs impact virus pathogenesis[193]. Functional studies have forged a link between viral miRNAs and the latent/lytic phase transition as well as host inflammatory response. This presents an interesting therapeutic option since these miRNAs are not in the human genome and interfering with their function should not have undesired side-effects.

Perhaps the most interesting example of a role for a miRNA in viral pathogenesis is the connection between miR-122 and hepatitis C virus (HCV). miR-122, highly expressed in hepatocytes, is required for efficient replication of the HCV virus[194]. Surprisingly,
miR-122 enhances replication HCV by promoting stability of the viral genomic RNA (in contrast to the conventional role of miRNAs in reducing target mRNA stability). Mutation of either miR-122 binding site in the HCV genome results in decreased viral RNA stability and virus replication[195, 196]. The reason HCV evolved to use a host cell miRNA is not clear, but may involve host cell trophism. A miR-122 inhibitor is undergoing clinical trials as an HCV therapeutic[197].

7. miRNAs as diagnostic markers

The extensive alterations in miRNA expression in disease provide great potential for clinical diagnostics based on miRNA signatures. Furthermore, miRNAs are more stable than mRNAs and can be recovered from formalin fixed paraffin sections (FFPE) and other sources with low overall RNA quality[198-200]. Current focus is on developing miRNA signatures for disease diagnosis, identifying cancers of unknown primary, and predicting response to therapy and drug resistance[201-208]. As yet, no in vitro diagnostic (IVD) using miRNA signatures has received FDA approval. However, a number of companies are offering Laboratory Developed Tests (LDT) that use miRNA signatures for cancer diagnostics[209, 210].

7.1. miRNA detection methods

miRNA detection methods can be divided into two general categories[211]. Discovery methods are designed to profile the expression of many miRNAs at once, often with high throughput. These approaches are largely centered on microarray hybridization and NGS profiling. The latter is especially powerful since it is possible to characterize novel miRNAs, while most other methods are limited to detection of known miRNA sequences. There are several NGS platforms available but all begin with the construction of template libraries. RNA adapters are ligated to both ends of small RNA fractions and the targets are RT-PCR amplified with primers directed at the adapters. This process amplifies all RNAs in the desired size range. The libraries can then be sequenced on several instruments, with the Illumina platforms probably the most common. Alternatively, single molecule instruments are available that avoid the PCR amplification step altogether[212, 213]. Since current instruments are capable of 200 million or more reads per library run, it is possible to multiplex 48 libraries (or more) in a single run and still achieve adequate sequence read depth[211]. Sequence reads are counted and quantitative expression profiles are obtained. As stated above, novel miRNAs and other small RNA species can be identified. While NGS based profiling has clear advantages and is becoming the default technology, nucleotide biases incurred during ligation steps has been observed[214, 215]. As with any profiling method a validation step is essential.

While NGS platforms are capable of high throughput profiling of the entire miRNA population, most clinical diagnostic approaches are based on rapid analysis of a small gene signature set. Therefore, RT-PCR and Nanostring are the focus of current diagnostic platforms[216]. Nanostring is a single molecule hybridization method that allows quantitiation of ~500 targets, either mRNA or miRNA, in a rapid experimental run[217]. An example of an LDT miRNA diagnostic is the pancreatic cancer test from Asuragen[218].
This RT-PCR based test uses a 7 miRNA signature to differentiate between pancreatic ductal adenocarcinoma and benign tissue.

7.2. miRNAs in extracellular fluids

Several years ago the startling discovery was made that miRNAs are present in human plasma[219, 220]. Due to the presence of RNase enzymes in plasma it was not expected that miRNAs would survive in this environment. It is now known that miRNAs are bound to protein, HDL, and/or encapsulated in lipid exosomes protecting them from degradation[221-227]. Originally thought to be cellular “debris” in plasma, possibly released from apoptotic cells, several studies have indicated that these extracellular miRNAs can be taken up by target cells and mediate target gene repression[228-230]. miRNA content in exosomes has distinct profiles from the parent cell, raising the possibility that miRNAs are sorted to exosomes in a controlled manner[231]. Thus, exosomes may function as extracellular messengers and may promote disease progression.

miRNAs have also been isolated from other extracellular fluids including saliva, urine, and stool[232]. Major efforts are underway to profile the miRNAs in these and other fluids, in normal subjects and in patients with disease[233]. Since many of these fluids can be easily obtained without invasive procedures they offer tremendous potential as a source material for clinical diagnostics. While numerous studies have reported extracellular miRNA signatures for diseases it remains to be determined whether these signatures have acceptable accuracy and sensitivity for the clinic.

8. miRNAs as therapeutic targets

As miRNAs become causally linked to increasing numbers of human diseases there are expanding efforts to develop therapeutics that directly target the miRNA[234-236]. These candidates are either miRNA mimics or antisense inhibitors, depending on the miRNA and underlying therapeutic goal. These drug candidates are based on modified nucleic acids and therefore present delivery challenges. The most promising candidate to date is the Regulus drug RG-101, currently in Phase II trials[197]. This drug is an antisense inhibitor of miR-122 that is being developed as a HCV therapeutic. A number of other miRNA therapeutic targets are being pursued by academia and industry, including targets for cardiovascular disease and cancer.

8.1. miRNA mimics and inhibitors

miRNA therapeutics fall under two general categories[234]. For miRNAs that are reduced in disease the therapeutic goal is to replace the miRNA. To achieve this a miRNA mimic is developed. These are typically dsRNA duplexes that directly load into RISC, though a stem-loop precursor can also be employed. In this case, the precursor is cleaved by Dicer and loaded into RISC. Since miRNA mimics are similar to siRNAs, the design strategies that have been developed for siRNA therapeutics have been adapted to miRNAs. Details of these designs are discussed in other sections of this special issue. Briefly, the duplex RNA requires backbone and ribose modifications to promote stability \textit{in vivo} and improve pharmacodynamics. Some strategy to promote cellular uptake must be included. This may
involve conjugation to sugar ligands or encapsulation in liposomes. Some therapeutic mimics that are under development are miR-34 for cancer (Phase I) and pre-clinical therapeutics including let-7 for cancer and miR-29 for fibrosis.[237, 238].

miRNAs that are elevated in disease can be targeted by antisense oligonucleotide inhibitors[234]. These inhibitors bind to the miRNA, most likely in the RISC complex, and prevent miRNA binding to target mRNAs. As with mimics, miRNA inhibitors require modifications for stability, and in some cases conjugation to ligands or encapsulation. The best example of a miRNA inhibitor is the RG101 anti-miR-122 in Phase II trials for HCV infection[197].

8.2. Small molecule inhibitors of the miRNA pathway

An interesting alternative to inhibiting the disease-associated miRNA directly is to develop small molecule compounds that modulate production or function of the miRNA[239]. This would avoid the problems associated with oligonucleotide delivery in vivo. Several research groups have performed high throughput screens for compounds that inhibit the function of specific miRNAs[240-245]. The readout for these screens is often a cell-based reporter that is normally silenced by the miRNA. If a compound interferes with production or activity of the miRNA, the reporter will be desilenced and the compound scored a hit. As an example, compounds that inhibit the activity of the oncogenic miR-21 have been identified using this approach[241, 242, 244]. The disadvantage of these small molecule drugs is that, by definition, the mechanism of action is unknown. There is much more downstream work required to define the specificity and direct mode of action.

9. Genome editing approaches

The miRNA pathway is a subset of post-transcriptional gene silencing (PTGS) pathways. In these related biological pathways the mRNA is the target of the silencing machinery, hence “post-transcriptional” silencing. A complementary approach to silence gene expression is to target the genome itself. Several approaches have been developed to redirect an endonuclease to a desired target site in genomic DNA[246]. The resultant double strand break is repaired, usually by non-homologous end joining, generating a short deletion at the repair site. This “editing” event is permanent and often leads to loss of expression of the targeted gene.

Early strategies for genome editing were based on the FokI nuclease domain fused to DNA binding protein modules[246]. Zinc finger nuclease (ZFN) and transcription activator like effector (TALE) domains are protein domains that bind to nucleic acids with predictable specificity. By constructing tandem arrangements of these domain variants it is possible to generate a DNA targeting protein with the desired sequence specificity. This targeting protein, fused to FokI, can direct permanent genome editing. The major factor that limited the widespread use of ZFNs and TALENs is the technical difficulty in construction of the targeting module. Multiple cloning steps are required, and since genome editing success is not guaranteed, many laboratories were dissuaded from the approach. This changed with the discovery of the CRISPR/Cas system.
The CRISPR/Cas pathway originates in bacterial where it functions as a primitive immune system[247]. Clustered regularly interspaced short palindromic repeats (CRISPR) are repeat sequences in bacterial genomes that are derived from pathogenic viruses. Bacteria use guide RNAs derived from these sequences to direct the Cas9 endonuclease to invading viruses, cutting the viral genome and preventing viral attack. By modifying the sequence of this guide RNA, it is possible to target any desired DNA sequence (see Figure 3)[248-251]. The resultant DNA break is repaired by NHEJ, leading to a short deletion. If a homologous DNA template is supplied, the break is repaired by homologous recombination, allowing precise control of the editing event. This approach has received rapid acceptance in molecular biological studies and has been used for disease modeling in cultured cells and in the mouse[252]. The major limitation for the use of CRISPR/Cas system as a therapeutic is that the Cas9 nuclease protein is bacterial in origin and must be ectopically expressed in the target cell. This requires the use of viral gene therapy vectors and the associated delivery and safety challenges that have limited clinical progress in that field. There is much excitement in the field, however, that CRISPR/Cas is a potential therapeutic approach, and studies in the mouse have demonstrated incredible potential[253].

10. Conclusion

We are entering an exciting time in the field of miRNA research. A decade of strong basic science research has forged links between specific miRNA alterations and many human diseases. We are now entering a phase where miRNA diagnostics and therapeutics can be developed with confidence, and several commercial efforts are underway to bring these developments to the clinic.

Acknowledgments

Work performed in the author's lab was supported by funding from the NIH (R01GM070674, RC1HL100108) and the Department of Cell Biology and Physiology.

List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RISC</td>
<td>RNA induced silencing complex</td>
</tr>
<tr>
<td>MCM7</td>
<td>Minichromosome maintenance protein 7</td>
</tr>
<tr>
<td>DGCR8</td>
<td>DiGeorge Syndrome critical region 8</td>
</tr>
<tr>
<td>TRBP</td>
<td>TAR RNA binding protein</td>
</tr>
<tr>
<td>NGS</td>
<td>Next generation sequencing</td>
</tr>
<tr>
<td>C3PO</td>
<td>Component 3 promoter of RISC</td>
</tr>
<tr>
<td>TNRC6</td>
<td>Trinucleotide repeat containing 6</td>
</tr>
<tr>
<td>CCR4-NOT</td>
<td>carbon catabolite repression 4 negative on TATA-less complex</td>
</tr>
<tr>
<td>tRF</td>
<td>tRNA derived fragment</td>
</tr>
<tr>
<td>tsRNA</td>
<td>tRNA derived small RNA</td>
</tr>
</tbody>
</table>
hnRNP  Heterogenous nuclear ribonuclear protein
C/EBP  CCAAT/enhancer binding protein
SRF  Serum response factor
MEF2  Myocyte enhancer factor 2
HCV  Hepatitis C virus
FDA  Food and Drug Administration
LDT  Laboratory developed test
FFPE  Formalin fixed paraffin embedded
IVD  In vitro diagnostic
siRNA  small interfering RNA
PTGS  Post-transcriptional gene silencing
ZFN  Zinc finger nuclease
TALEN  Transcription activator like effector nuclease
CRISPR  Clustered regularly interspaced short palindromic repeats
NHEJ  Non-homologous end joining
PAM  Protoscaler adjacent motif

References


Adv Drug Deliv Rev. Author manuscript; available in PMC 2016 June 29.


209. Asuragen.
210. RosettaGenomics.


Figure 1. The miRNA biogenesis pathway
The canonical miRNA biogenesis pathway is shown in the central part of the figure. miRNAs are transcribed by RNA polymerase II. The pri-miRNA transcript is cut by the enzyme Drosha yielding the precursor miRNA. After Dicer cleavage the mature miRNA is loaded into the effector complex RISC where it directs translational repression of mRNA targets. Two well characterized non-canonical biogenesis routes are shown in the semi-transparent graphic. mirtrons bypass the Drosha processing step; a splicing reaction produces the precursor. Subsequent biogenesis steps are the same. The second non-canonical route shown is specific for miR-451. This miRNA precursor is directly bound and cut by
Argonaute 2 in a Dicer independent reaction. The remaining RNA is trimmed back to yield a mature miRNA.
Figure 2. Anatomy of a miRNA poly-cistron
The miRNA cluster containing miR-106b, miR-93, and miR-25 is shown in its genomic context. Three miRNAs are located within an intron of the protein coding gene MCM7. Boxes in the MCM7 gene model represent exons, and the arrow line represents introns. Sequence conservation is shown. Data were adapted from the UCSC genome browser.
Figure 3. CRISPR/Cas as a genome editing tool
The guide RNA directs the Cas9 endonuclease to target DNA. Modification of the guide RNA sequence allows any DNA sequence to be targeted. The only sequence requirement is the PAM motif NCC downstream of the target site, and a C at the 5’ end of the target site. The double strand break is repaired by non-homologous end joining, leading to a short deletion.