How Do MicroRNAs Regulate Gene Expression?

Richard J. Jackson* and Nancy Standart*

(Published 2 January 2007)

Several thousand human genes, amounting to about one-third of the whole genome, are potential targets for regulation by the several hundred microRNAs (miRNAs) encoded in the genome. The regulation occurs posttranscriptionally and involves the ~21-nucleotide miRNA interacting with a target site in the mRNA that generally has imperfect complementarity to the miRNA. The target sites are almost invariably in the 3′-untranslated region of the messenger RNA (mRNA), often in multiple copies. Metazoan miRNAs were previously thought to down-regulate protein expression by inhibiting target mRNA translation at some stage after the translation initiation step, without much effect on mRNA abundance. However, recent studies have questioned these suppositions. With some targets, an increase in the rate of mRNA degradation by the normal decay pathway contributes to the decrease in protein expression. miRNAs can also inhibit translation initiation, specifically the function of the cap-binding initiation factor, eIF4E. Repressed target mRNAs as well as miRNAs themselves accumulate in cytoplasmic foci known as P-bodies, where many enzymes involved in mRNA degradation are concentrated. However, P-bodies may also serve as repositories for the temporary and reversible storage of untranslated mRNA, and reducing the expression (knockdown) of several distinct P-body protein components can alleviate miRNA-mediated repression of gene expression.

Introduction

A PubMed search of "microRNA" or "miRNA" gives just over 1200 citations, all of them published since the turn of the millennium. These numerous publications have been concerned primarily with using experimental and in silico approaches to identify miRNAs and the target genes that they regulate in various organisms, elucidating the biogenesis of miRNAs, and investigating how the spectrum of miRNAs differs in different tissues and disease states. Thus, by mid-2005, it was known that the human genome encodes several hundred different microRNAs (*1*) and that the pattern of miRNA expression is often perturbed in disease states (*2*–*4*). Bioinformatic approaches (*1*) further suggested that the mammalian miRNA repertoire might collectively regulate several thousand genes, even though only a handful of these predicted targets have been validated to date. By contrast, very few of these studies addressed the mechanism of miRNA-mediated repression. Primary publications and reviews written before that date generally proposed that miRNAs (i) do not promote degradation of their target mRNAs, but (ii) they do down-regulate target mRNA translation at some stage after the initiation step.

Over the past 18 months, there has been a steady flow of provocative reports on mechanisms of miRNA-mediated repression, and, far from consolidating and building upon these earlier ideas, many of them directly contradict one or another of these original assertions. The aim of this Review is to illuminate and explain the controversies generated by these recent publications. However, we have been unable to resolve many of the apparent contradictions, and our advice to the reader is to keep an open mind to the possibility that there may be more than one mechanism by which miRNAs effect posttranscriptional regulation of gene expression.

This Review will therefore focus specifically on mechanisms, with particular reference to metazoans. Lower eukaryotes are mentioned only to the extent that they have contributed to our understanding of mechanisms likely to operate in higher organisms, and plants are ignored altogether, as the mechanisms of plant miRNA biogenesis and action seem distinctly different. Mechanisms of RNA interference (RNAi) by short interfering RNAs (siRNAs) are also beyond the scope of this Review except insofar as they relate directly to miRNA-mediated repression.

siRNAs Versus miRNAs: Definitions, Similarities, and Differences

The generally accepted distinction between "miRNAs" and "siRNAs" is that the former are ~21-residue RNAs derived from longer RNAs that include a ~70-nucleotide (nt) imperfectly base-paired hairpin segment that is the precursor of the mature miRNA, whereas siRNAs are of similar length but are derived from longer, perfectly complementary double-stranded RNA precursors of endogenous or exogenous origin (*5*). In both cases, biogenesis proceeds through a staggered duplex intermediate, usually with ~19 base pairs (bp), and always with unpaired 3′ extensions of 2 nt and a 5′-phosphate on each strand. Especially in the case of mammalian cells, these ~21-nt RNAs are often introduced by transfection of chemically synthesized versions of this staggered duplex RNA, or alternatively as DNA constructs that will give rise to such a duplex after transcription and processing. One of the two strands, the one with the lower-stability base-pairing at its 5′ end, is then assimilated into a complex with several proteins, while the other, the so-called passenger strand, is degraded [reviewed in (*6*)]. The resulting RNA-protein complex is generally known as RISC (RNA-induced silencing complex), or for miRNAs, it is sometimes known as an miRNP (miRNA-protein complex).

For RNAi, the 21-residue siRNA, or its longer double-stranded RNA precursor, is designed to be perfectly complementary to part of the target mRNA sequence, which leads to endonucleolytic cleavage of the mRNA at the site of complementarity (*6*),

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK.

***E-mail: rjj@mole.bio.cam.ac.uk (R.J.J.); nms@mole.bio.cam.ac.uk (N.S.)**

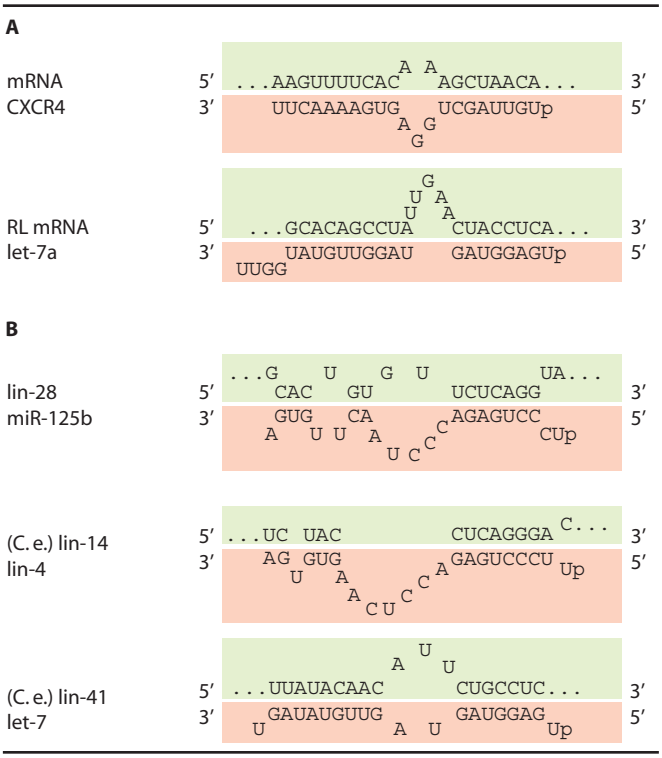

Fig. 1. Examples of the imperfect complementarity between various miRNAs and their bulged target sites. Micro-RNA sequences are in red, mRNA target sites in green. The two miRNA/mRNA pairs in (**A**) have been widely used in studies of the mechanisms of miRNA-mediated repression discussed here. The first is an artificial system based on an siRNA designed for RNAi of CXCR4 mRNA, but adapted to serve as a miRNA mimic through the design of bulged target sites in the reporter mRNA 3′-UTR. The second is the interaction of (endogenous) let-7a miRNA with a reporter target site based on the predicted interaction of let-7a with human *lin-28* mRNA. The three examples in (**B**) are endogenous miRNAs and their predicted (miR-125b/human *lin-28* mRNA) or validated (*C. elegans lin-4* and *let-7* miRNAs) targets sites. Relevant references cited in this work are as follows (note that these are not necessarily the publications that first identified the miRNA/target site interaction): mRNA/CXCR4 (*8*, *11*, *12*, *33*, *79*); RL mRNA/let-7a (*29*); *lin-28*/miR-125b (*47*); (C.e.) *lin-14/lin-4* (*22*–*25*, *49*); (C.e.) *lin-41/let-7* (*49*).

followed by degradation of the two fragments by enzymes of the normal mRNA decay pathways. Cleavage may not be entirely abrogated by a very limited number of mismatches, depending on their position, and so miRNAs that have a similar near-perfect complementarity to their target sites can also promote degradation by this route (*7*). Generally, however, the interaction of an miRNA with its predicted or validated target involves quite extensive mismatched bulges, especially in the central region (Fig. 1), and to a lesser extent in the miRNA 3′ end. This type of interaction does not generally lead to mRNA degradation via endonucleolytic cleavage, but results in a decrease in protein expression that is usually (but not always) greater than the decrease in mRNA abundance, suggesting that the main reason for down-regulation is reduced efficiency of translation rather than increased mRNA degradation. Validated and predicted miRNA target sites are almost invariably located in the 3′-untranslated region (3′-UTR) of a gene, often in multiple copies.

It is important to note that the formal distinction between an siRNA and an miRNA is based solely on their different origins and biogenesis, irrespective of the degree of complementarity to the target site and irrespective of whether the ~21-nt RNA promotes endonucleolytic cleavage of the target mRNA, or repression of the target by some other mechanism. According to these definitions, a ~21-nt RNA introduced into cells by transfecting a chemically synthesized staggered duplex intermediate should be called an "siRNA," even if its sequence has been designed to be precisely the same as that of a natural miRNA. This is a potential source of confusion to those outside the immediate field, and in order to avoid such difficulties, it might be preferable to refer to such an RNA as a si(mi)RNA, to highlight the fact that although it is technically an siRNA, in fact it is one that is designed to exactly mimic a bona fide miRNA.

Because the different outcomes depend on the degree of complementarity between the short RNA and its target, it follows that a ~21-nt RNA designed to promote endonucleolytic cleavage of a particular mRNA with a perfectly complementary target site could also act as a translational repressor of another mRNA if this has appropriate mismatched bulged target sites. The validity of this argument has been demonstrated for an siRNA designed for RNAi of mRNA encoding the CXCR4 chemokine receptor (*8*). This siRNA has subsequently been widely used to study how interaction of a ~21-nt RNA with imperfectly complementary 3′-UTR sites results in repression of mRNA translation. For the reasons explained above and because the sequence of this small RNA shows no resemblance to any known natural miRNA, we will refer to it as CXCR4 siRNA, but, except where otherwise stated, it should be understood that it is always being used in conjunction with mismatched target sites and is therefore serving as a surrogate for a typical microRNA.

The mismatch between an miRNA and its target site can take on different configurations, with a central unpaired bulge in either the miRNA or mRNA strand, or both. Figure 1 shows both natural miRNAs paired to their natural target, and the pairing of the artificial CXCR4 siRNA with its laboratory-designed target site. In the latter case, the exact configuration of the unpaired bulge(s) affects the degree of repression, but this may not be applicable for other miRNA-mRNA pairs (*9*). Two extensive studies of how target site sequence and complementarity influence repression by either natural *Drosophila* miRNAs overexpressed in the wing imaginal disc or by the CXCR4 siRNA transfected into HeLa cells are in very close agreement as to what is important (*10*, *11*). First and foremost, residues in the 5′ portion of the miRNA (residues 2 to 8, the so-called seed) upstream of the discontinuity should be (almost) perfectly complementary to the mRNA. The degree of repression is dependent on the stability of the pairing in this region, with a rather sharp cut-off, and in addition, A-form geometry seems important because G-U pairs in this region decrease repression to a greater extent than can be explained purely on the grounds of mRNA stability. If the 5′ portion of the miRNA is optimally base-paired, pairing between the 3′ portion and the mRNA target is not critical, but a high degree of complementarity here can rescue repression when the pairing of the miRNA "seed" is marginally suboptimal. In the CXCR4 system, the degree of repression is related to the number of target sites, at least up to six such sites, the maximum tested so far (*8*, *12*). In addition, different miRNAs can act in a combinatorial

way, in the sense that cotransfection of two different siRNAs with a reporter construct that had two 3′-UTR bulged target sites for each of them resulted in a similar degree of repression, as when there were four identical target sites for just one of the siRNAs (*11*). Many endogenous mRNAs have just one or two predicted sites for interaction with endogenous miRNAs, yet seem to be quite efficiently repressed, which raises questions as to whether the artificial systems such as CXCR4 may be intrinsically less potent repressors in some way. This idea is given some support by recent indications that the sequences flanking miRNA target sites, or the context of such sites, can influence the biological outcome of miRNA-mRNA interactions (*13*).

The contrasting outcomes that are dependent on whether interaction of the ~21-nt small RNA with the mRNA target site is perfectly complementary or has substantial bulged mismatches are presumably due to different proteins being present in the RISC and miRNP or, if the protein composition is really identical, then to the degree of complementarity influencing how these proteins act. The protein compositions of RISC assembled on perfectly complementary siRNAs of exogenous origin, and the miRNPs formed with endogenous miRNAs, are often described as similar or overlapping rather than absolutely identical. However, as the functional distinction between between a perfectly complementary interaction and one with mismatches is only apparent after the ~21-nt RNA has interacted with the target site, any differences are likely to arise concomitantly with, or shortly after, the establishment of this interaction.

Both types of complex include members of the Argonaute family of proteins. Of the four mammalian Argonaute proteins, only Ago2 is capable of siRNA-mediated endonucleolytic cleavage, with its ribonuclease H (RNase H)–like domain cutting the phosphodiester bond in the target mRNA opposite the 10th and 11th residue in the siRNA (*14*, *15*). Cocrystal structures show that the 5′-phosphate on the siRNA is essential to place this phosphodiester bond in the active site (*16*, *17*). Such endonucleolytic cleavage would not be expected with most miRNA-mRNA interactions because the A-form geometry would be severely disrupted by the unpaired bulges in this region (Fig. 1).

Knockdown (via RNAi) of each Ago protein individually in human embryonic kidney (HEK) 293 cells showed that Ago2 was the major contributor to miRNA-mediated repression of target mRNA translation (*18*). Nevertheless, there is evidence that Ago1, Ago3, and Ago4 are each capable of promoting this repression to some extent (*14*, *19*), even though they have no known enzymatic activity. In contrast, the two Argonaute proteins in *Drosophila* appear to have largely nonoverlapping functions, with Ago1 involved mainly or exclusively in miRNA-mediated repression, and Ago2 restricted to siRNA-mediated endonucleolytic cleavage (*20*, *21*).

Evidence That miRNAs Can Inhibit Translation at Some Stage After Initiation

The pre-2005 hypothesis that miRNA-mediated repression of target gene expression is due to inhibition of mRNA translation at some stage after the initiation step was based on just two studies of a single *Caenorhabditis elegans* miRNA, *lin-4* miRNA (the first microRNA to be discovered, long before the designation "microRNA" had been invented), and two larval mRNA targets, *lin-14* mRNA, which has up to seven potential 3′-UTR sites for bulged interaction with *lin-4* miRNA (*22*, *23*), and *lin-28* mRNA. Expression of *lin-4* miRNA commences late in the L1 larval stage, and by the late L2 or early L3 stage, the amounts of LIN-14 and LIN-28 proteins are less than 10% of that seen in mid-L1, but the abundance of *lin-28* mRNA is essentially unchanged, whereas that of *lin-14* mRNA is about half of that in the L1 stage (*24*, *25*), with no appreciable difference in poly(A) tail length (15 to 30 A residues).

In sucrose gradient analyses of polysomes at the late L2 or early L3 stage, the repressed *lin-14* and *lin-28* mRNAs were found in the same fractions as in a polysome distribution analysis of mid-L1 larvae, where both would be efficiently translated (*24*, *25*). When these polysomes from late L2 larvae were analyzed on gradients containing EDTA, both mRNAs were found near the top of the gradient, and on metrizamide gradients the *lin-14* mRNA was found in fractions of the same buoyant density as those containing polysomes from L1 larvae (*24*, *25*). Thus, by these two commonly used criteria, *lin-14* and *lin-28* mRNAs appeared to be genuinely in polysomes in late L2 and early L3 larvae, even though their translation was repressed at these stages. These tests do not, however, prove that the polysomes are dynamic and capable of elongation, an issue that can be examined by seeing whether the repressed mRNA moves into smaller polysomes on incubation with puromycin, or with a specific inhibitor of initiation (polysome runoff assays). These functional tests are best done in intact cells, in case the translational repression might be relieved by the mere act of preparing a cell-free extract, and should be done over a relatively short time period (~5 min) so that differences in the rate of reaction with puromycin or the rate of runoff can be detected. However, puromycin and specific initiation inhibitors do not penetrate *C. elegans* larvae, and so the best that could be done was to isolate polysomes from early L3 larvae and add them to reticulocyte lysate for polysome runoff, which showed that in a rather long incubation period (45 min), the *lin-28* mRNA moved out of polysomes as effectively as did an mRNA that was not repressed in the L3 larvae (*25*). The question of whether this runoff produced any full-length LIN-28 protein was not examined.

The results described above are not a peculiarity of *C. elegans* larval development. Similar results were obtained when DNA transfections of 293T cells were used to study the repression of a luciferase reporter with six bulged sites for interaction with CXCR4 siRNA. Introduction of the CXCR4 siRNA reduced overall luciferase expression by more than 95%, with relatively little effect on target mRNA abundance (*12*). Likewise, the presence or absence of the CXCR4 siRNA had no influence on the polysomal distribution of the target mRNA, nor did it affect the movement of this mRNA into smaller polysomes on a brief (3 min) incubation of the cells with puromycin. On a 5-min incubation with a specific inhibitor of initiation, the target mRNA moved into smaller polysomes, as expected, but the shift was distinctly greater under conditions of repression by CXCR4 siRNA. This was interpreted as premature ribosome drop-off, or a failure of processivity (*12*). However, if such drop-off was occurring continuously, it is hard to see why the steady-state polysomal distribution of the target mRNA under repressed conditions would be so similar to the unrepressed distribution, unless there was a fortuitous increase in initiation frequency that precisely compensated for the decrease in ribosome transit time. A plausible alternative explanation is that the target mRNA is slightly more sensitive to the initiation inhibitor when it is repressed by the CXCR4 siRNA than it is under unrepressed conditions.

In this system, both cistrons of a dicistronic reporter with either the hepatitis C virus (HCV) or cricket paralysis virus