RNA interference

Literature: to read one or the other of these reviews is highly recommended

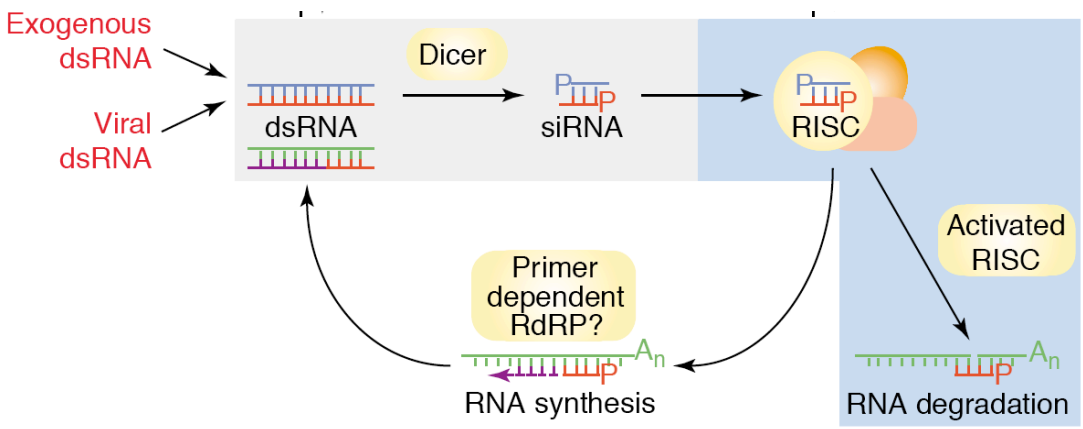
1. Cullen, B. R. (2002). RNA interference: antiviral defense and genetic tool, Nature Immunol *3*, 597-9.
2. Hannon, G. J. (2002). RNA interference, Nature *418*, 244-51.
3. McManus, M. T., and Sharp, P. A. (2002). Gene silencing in mammals by small interfering RNAs, Nature Reviews Genetics *3*, 737-47.
4. Meister G and Tuschl T (2004) Mechanisms of gene silencing by double-stranded RNA, Nature 431, 343-349
5. Special Issue of FBES letters: RNAi mechanisms, biology and applications 31 October 2005 (a collection of excellent reviews: eg the ones by SM Hammond or Wienholds and Plasterk)
6. **Ha and Kim (2014) Regulation of miRNA biogenesis. Nature Reviews Molecular Cell Biology 15, 509 – 524.**
7. Stefani, G. and Slack, F.J. (2008) Small non-coding RNAs in animal development. Nature Reviews Molecular Cell Biology 9, 219-230
8. Senti and Brennecke (2010) The piRNA pathway: a fly's perspective… Trends in Genetics 26(12), 499-509
9. **Jonas and Izaurralde (2105) Towards a molecular understanding of microRNA-mediated gene silencing. Nature Reviews Genetics 16, 421-433.**

1. Factors and mechanism

RNA interference (RNAi) as a protecting mechanism against invasion by foreign genes and has turned out to be an invaluable technical tool to interfere with gene function in higher eukaryotes. It was first observed in plants and then described in C.elegans and has subsequently been demonstrated in diverse eukaryotes such as insects, fungi and vertebrates.

RNAi is the mechanism of sequence-specific, post-transcriptional gene silencing initiated by double-stranded RNAs (dsRNA) or microRNAs (miRNAs) homologous to the gene being suppressed. However, the RNAi machinery may also affect gene expression at the level chromatin structure in various organisms.

GENERAL MODEL



dsRNAs are processed by Dicer, a cellular ribonuclease III, to generate duplexes of about 21 nt with 3'-overhangs (small interfering RNA, siRNA) which mediate sequence-specific mRNA degradation. The exact length of these RNA duplexes is species specific and caused by differences in the positioning of the catalytic center of Dicer on the dsRNA substrate.

DICER

Dicer is an RNase III-related endonuclease responsible for processing double-stranded RNA (dsRNA) to small interfering RNAs (siRNAs) and the excision of the regulatory microRNAs from their precursors.

Dicer is a large (220 kDa) multi-domain protein present in all eukaryotes studied to date, with the exception of budding yeast. Its domains include a putative RNA helicase/ ATPase domain, a PAZ domain, two neighbouring RNase III-like domains and a dsRNA-binding domain (RBD). Dicer processes precursor dsRNAs to make both siRNAs and miRNAs (see below).

RISC

In the effector step, the double-stranded siRNAs are incorporated into a multiprotein complex, known **as the RNA-induced silencing complex (RISC)**. RISC is then proposed to undergo an ATP-dependent activation step that results in the unwinding of the double-stranded siRNAs. Activated RISC uses a single-stranded siRNA as a guide to identify complementary RNAs and an endoribonuclease (a member of the Argonaute protein family) then cleaves the target RNA across from the center of the guide siRNA. Finally, the cleaved RNA is probably degraded by exoribonucleases.

RdRP

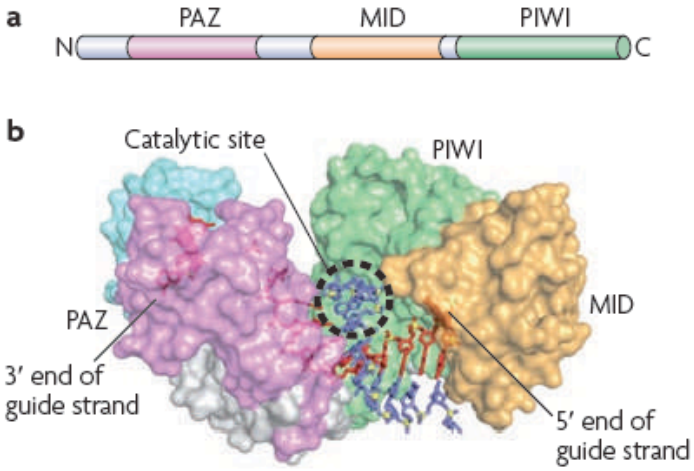
Genetic analyses suggest that an amplification step might be required for efficient RNA-mediated silencing in several systems. Recent studies in C. elegans have led to a model in which primary siRNAs (derived from the trigger dsRNA) might prime the synthesis of additional dsRNA, using the target mRNA as a template, in a reaction catalyzed by an **RNA-directed RNA polymerase (RdRP)**. The newly synthesized dsRNA would then be cleaved by Dicer to generate secondary siRNAs at a sufficient concentration to achieve efficient target mRNA degradation by RISC.

Argonaute proteins

Member of the Argonaute gene family have been implicated in RNA-mediated silencing in several eukaryotes and many of them are found in association with RISC. Argonaute proteins are characterized by the presence of two structural domains: a PAZ (for `Piwi/Argonaute/Zwille') domain and a C-terminal Piwi domain. Recently, it has been shown that the Piwi domain resembles RNase H. Biochemical evidence suggests that Argonaute 2 is the enzyme ’slicer’, which mediates mRNA cleavage as part of the RISC complex.

Box 1 | **Argonaute proteins and their associated small RNAs**

The Argonaute (Ago) family can be classified into two subclades: the Ago subfamily and the Piwi subfamily (TABLE 1). The Ago proteins are expressed ubiquitously, interact with microRNAs (miRNAs) or small interfering RNAs (siRNAs), and function as post-transcriptional regulators. The Piwi proteins are abundantly expressed in germ cells and function in transposon silencing, together with Piwi-interacting RNAs (piRNAs).



Ago-family proteins are composed of three characteristic domains: the PAZ, MID and PIWI domains (see the figure, part **a**). The PAZ domain serves as a docking site for the 3′ end of small RNA[181–184], whereas the MID domain anchors the 5′ terminal nucleotide[183–187] (see the figure, part **b**). Recent studies have determined the structure of *Thermus thermophilus* Ago with a guide strand and target strand duplex[183,184]. The PIWI domain has a structure that is similar to RNase H, which cuts the RNA strand of an RNA–DNA hybrid. Indeed, the PIWI domain of some Ago proteins can cleave the target RNA bound to small RNA: this is called slicer activity. Of the four human Ago proteins (AGO1–4; also known as EIF2C1–4), only AGO2 has slicer activity, whereas in *Drosophila melanogaster* all Ago and Piwi proteins possess slicer activity. Apart from the endonucleolytic cleavage that is mediated by the PIWI domain, the Ago proteins can induce translational repression and exonucleolytic mRNA decay through interaction with other protein factors[13].

Micro RNAs

MicroRNAs constitute a class of non-coding small RNAs that are phylogenetically widespread in invertebrates, vertebrates and plants. The small temporal RNA (stRNAs) lin-4 and let-7, which belong to a subclass of miRNAs, were initially identified in *C.elegans* as essential regulators of the timing of development.

In animals, precursor miRNAs (pre-miRNAs) are transcribed a primary microRNAs (pri-miRNA) that are cleaved in the nucleus by the RNaseIII family member Drosha to pre-miRNAs of 60–70 nt in length. The pre-miRNAs fold into imperfectly pairing inverted repeats forming stem loop structures. After their export to the cytoplasm by Exportin5, these pre-miRNAs are processed to mature miRNAs by DICER.

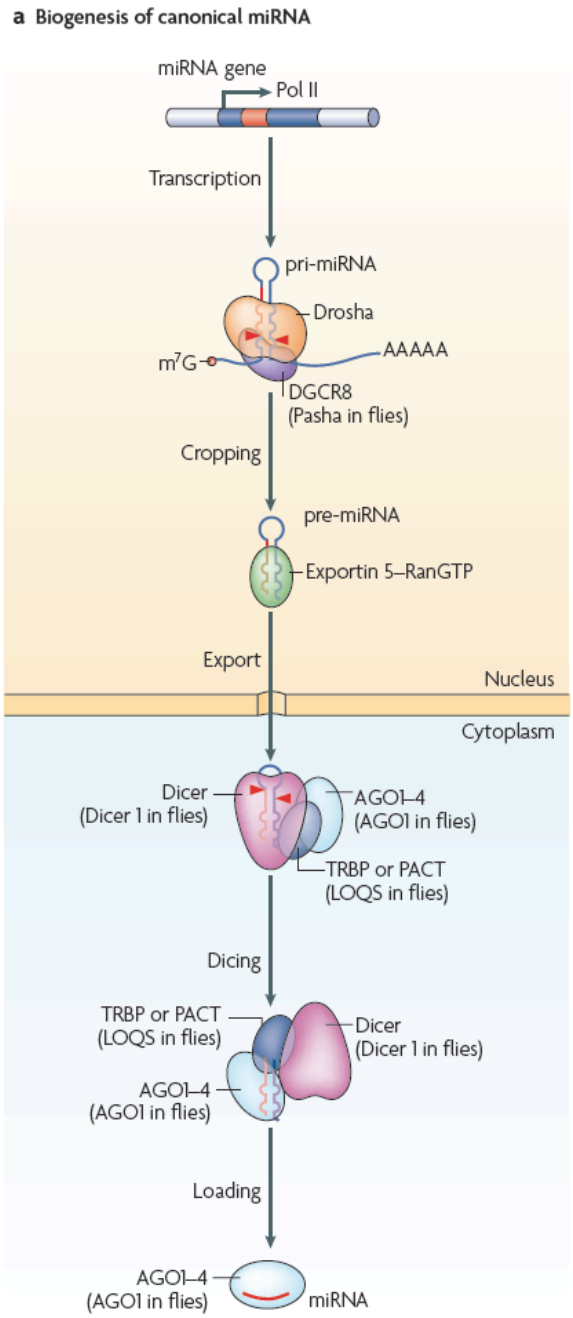
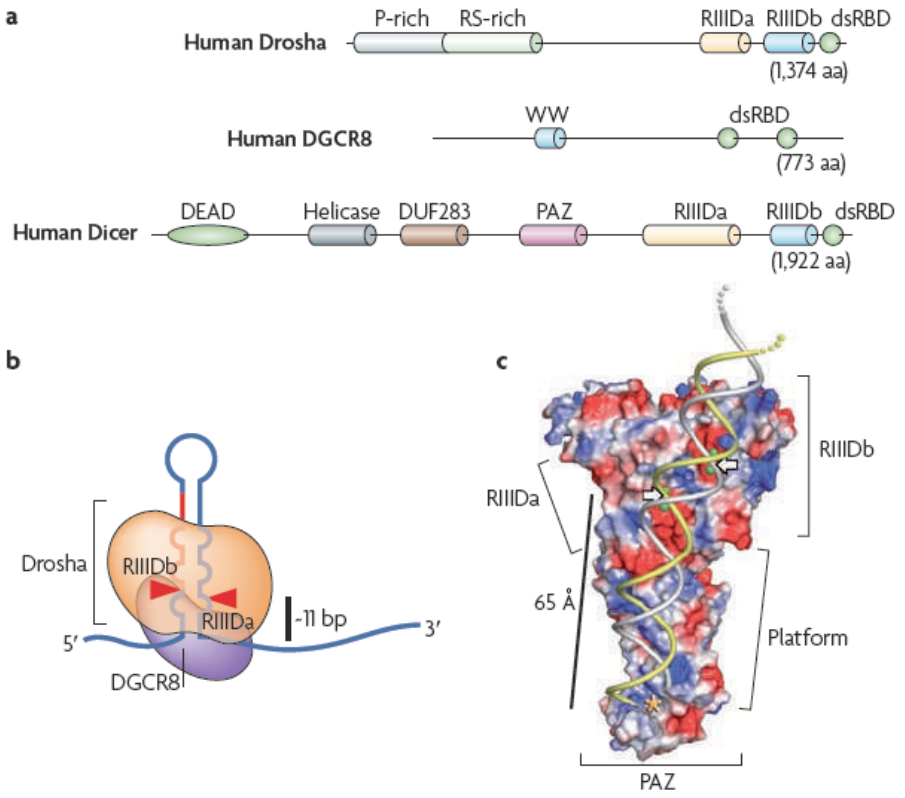


Figure 2 | **miRNA biogenesis pathway. a** | Canonical microRNA (miRNA) genes are transcribed by RNA polymerase II (Pol II) to generate the primary transcripts (pri-miRNAs). The initiation step (cropping) is mediated by the Drosha–DiGeorge syndrome critical region gene 8 (DGCR8; Pasha in *Drosophila melanogaster* and *Caenorhabditis elegans*) complex (also known as the Microprocessor complex) that generates ~65 nucleotide (nt) pre-miRNAs. Pre-miRNA has a short stem plus a ~2-nt 3′ overhang, which is recognized by the nuclear export factor exportin 5 (EXP5). On export from the nucleus, the cytoplasmic RNase III Dicer catalyses the second processing (dicing) step to produce miRNA duplexes. Dicer, TRBP (TAR RNA-binding protein; also known as TARBP2) or PACT (also known as PRKRA), and Argonaute (AGO)1–4 (also known as EIF2C1–4) mediate the processing of pre-miRNA and the assembly of the RISC (RNA-induced silencing complex) in humans. One strand of the duplex remains on the Ago protein as the mature miRNA, whereas the other strand is degraded. Ago is thought to be associated with Dicer in the dicing step as well as in the RISC assembly step. In *D. melanogaster*, Dicer 1, Loquacious (LOQS; also known as R3D1) and AGO1 are responsible for the same process. In flies, most miRNAs are loaded onto AGO1, whereas miRNAs from highly base-paired precursors are sorted into AGO2. The figure shows the mammalian processing pathways with fly components in brackets. **b** | Canonical

Box 2 | **RNase III proteins and their mechanism of action**



Two types of RNase III are found in animals: Drosha and Dicer. Both proteins have two tandem RNase III domains (RIIIDs) and a double-stranded RNA-binding domain (dsRBD; see the figure, part **a**). Two RIIIDs interact with each other to make an intramolecular dimer in which the two catalytic sites are located closely to each other. The first RIIID cuts the 3′ strand of dsRNA, whereas the second RIIID cleaves the 5′ strand, generating a 2-nucleotide (nt) 3′ overhang[38,189,190] (see the figure, parts **b**,**c**).

Drosha is a nuclear protein of 130–160 kDa. The dsRBD is necessary for activity, although it is not sufficient to bind a primary transcript (pri-miRNA). To provide the RNA-binding activity, Drosha interacts with a dsRNA-binding protein, DiGeorge syndrome critical region gene 8 (DGCR8; Pasha in *Drosophila melanogaster* and *Caenorhabditis elegans*), through its middle region (this Drosha–DGCR8 complex is known as the Microprocessor; see the figure, part **b**). The amino-terminal region contains the nuclear localization signal, which is dispensable for cleavage activity *in vitro*[38,191].

DGCR8 and Pasha are ~120 kDa and localize to the nucleoplasm and the nucleolus[177,191]. They contain two dsRBDs that recognize the ssRNA–dsRNA junction of a substrate[46]. DGCR8 and Pasha interact with Drosha through their C-terminal domain[191]. A structure of the partial human DGCR8 protein that contains the two dsRBDs and a part of the carboxy-terminal domain showed that the two dsRBDs are arranged with pseudo two-fold symmetry and that the C-terminal helix is closely packed against the two dsRBDs[192]. The middle domain of DGCR8 is required for haem binding, oligomerization and pri-miRNA processing[191,193].

Dicer homologues are cytoplasmic RNase III proteins of ~200 kDa. The middle region of Dicer contains a PAZ domain, which binds to the 3′ protruding end of RNAs[181,182,194,195]. Structural studies on the *Giardia intestinalis* Dicer protein show that the processing centre made of the two RIIIDs is connected to the PAZ domain by a long, positively charged helix[196]. Based on the structure (see the figure, part **c**), it is predicted that the 3′ end of a dsRNA is docked to the PAZ domain and the RNA stem interacts with the flat, positively charged extension to reach the catalytic centre. The distance between the PAZ domain and the catalytic site (65 Å) approximately matches the length of the product of *G. intestinalis* Dicer (25 nt). The DEAD-box RNA helicase domain is not necessary for Dicer activity *in vitro*[78], and its role remains unknown. Figure part **c** is reproduced, with permission, from REF. 196 © (2006) American Association for the Advancement of Science.

Mature miRNAs are about 21 nt and can form imperfect duplexes with sequences in the 3' untranslated regions of target mRNAs, leading to translational repression and mRNA degradation. The AGO-associated factor GW182 plays a pivotal role triggering mRNA deadenylation and decapping.
In some cases (e.g. in plants) the miRNAs form perfect duplexes with target mRNAs in the cytoplasm (and induce mRNA degradation).

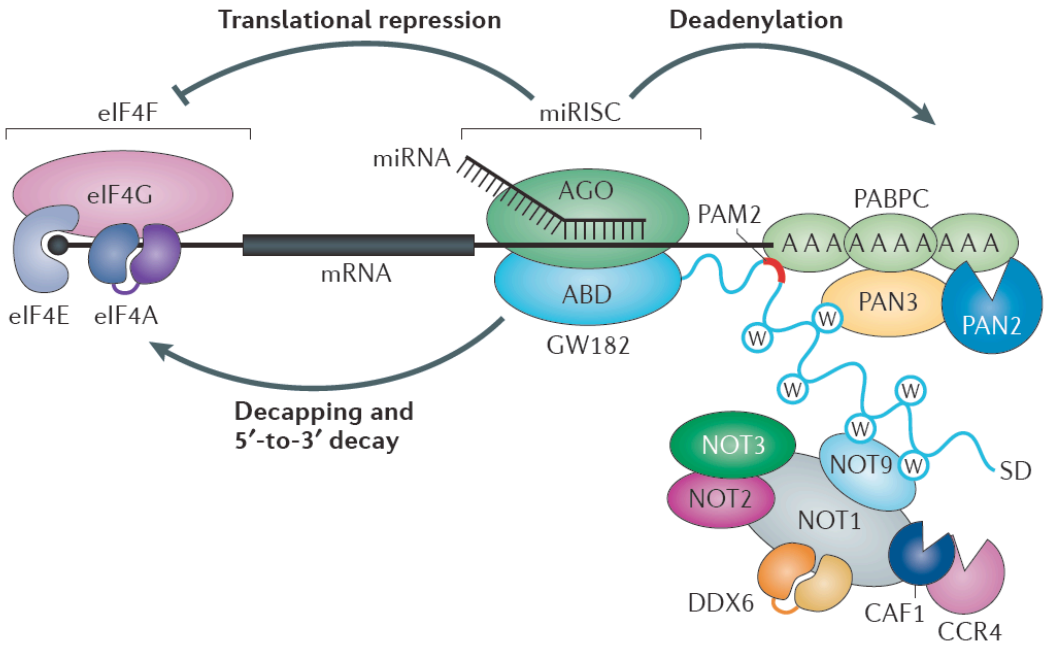


Figure 1 | **Overview of miRNA-mediated gene silencing in animals.** Animal microRNAs (miRNAs) bound to an Argonaute (AGO) protein in miRNA-induced silencing complexes (miRISCs) recognize their mRNA targets by base-pairing to partially complementary binding sites, which are predominantly located in the 3′ untranslated region of the mRNA. AGO proteins interact with a GW182 protein, which in turn interacts with cytoplasmic poly(A)-binding protein (PABPC) and with the cytoplasmic deadenylase complexes PAN2–PAN3 and CCR4–NOT. The PAN2–PAN3 and CCR4–NOT complexes catalyse the deadenylation of the mRNA target. The GW182 proteins consist of an amino-terminal AGO-binding domain (ABD) and a silencing domain (SD; see BOX 2 for additional information). In animal cell cultures, deadenylated mRNAs are decapped and rapidly degraded by 5′-to-3′ exoribonuclease 1 (XRN1; not shown) (BOX 1). In addition, miRNAs repress translation, but the precise molecular mechanism for this remains unclear. The emerging consensus is that miRNAs inhibit translation initiation by interfering with the activity and/or assembly of the eukaryotic initiation factor 4F (eIF4F) complex. The eIF4F complex consists of the cap-binding protein eIF4E, the adaptor protein eIF4G and the DEAD box RNA helicase eIF4A[99]. eIF4G serves as a scaffold for protein–protein interactions that are essential for the recruitment of the 43S pre-initiation complex and for translation initiation[99]. The cap structure is shown as a black circle. DDX6, DEAD box protein 6; PAM2, PABP-interacting motif 2.

RNAi and influences on chromatin

In several plant species, dsRNA can direct methylation of homologous DNA sequences. Methylation of genomic DNA occurs even when silencing is induced by RNA viruses bearing sequences homologous to nuclear DNA that replicate exclusively in the cytoplasm. This suggests that there is communication between the cytoplasm and the nucleus transferring the information. When the dsRNA has homology to a promoter, it induces transcriptional silencing in association with DNA methylation, however, it is not known whether long dsRNA or processed small RNAs are involved in this process.
Connections between the RNAi and the PTGS machinery and chromatin and/or genomic DNA modifications are also starting to emerge in other organisms.

Several recent reports have directly implicated the RNAi and PTGS machinery in heterochromatin formation and genome rearrangements. In many eukaryotes, heterochromatin is characterized by a high density of histone H3 methylated at lysine 9 (H3-Lys9). This modification results in the binding to histone H3 of heterochromatin protein 1 (HP1), and presumably other factors, and formation of a transcriptionally repressive chromatin structure. But DNA methylation might be a secondary modification that contributes, in certain organisms, to the stability and inheritability of the silent chromatin state.

Although small RNAs presumably target chromatin modification through a pairing mechanism, the recognition step and the components that link small RNAs to histone modification are currently unclear. Nevertheless, these observations extend the range of dsRNA-mediated processes and encourage the examination of transcriptional regulation by dsRNA.

Biological functions of RNA-mediated silencing

Mutational inactivation of components of the RNAi and PTGS machinery affects at least three distinct eukaryotic processes:

- defense response against viruses
- transposon mobility and
- regulation of development in multicellular organisms

The RNAi and PTGS processes were originally proposed to have evolved to counteract genomic parasites, but it is becoming apparent that dsRNA-mediated mechanisms are also involved in the normal regulation of endogenous genes.

Intriguingly, a high proportion of the predicted miRNA targets function as developmental regulators, suggesting that miRNAs might have a role in coordinating growth and development. In C. elegans, developmental defects resulting from reduced function of Dicer and the Argonaute-like proteins ALG-1 and ALG-2 have also been attributed to the improper processing of miRNA

precursors and a reduction in mature stRNA expression. It is therefore tempting to propose the existence of ancient, miRNA-mediated mechanisms that regulate endogenous genes in eukaryotes.

It is clear that despite much progress resulting from a combination of genetics and biochemistry, we are only just beginning to understand the mechanistic complexity of RNA-mediated silencing, its biological implications, and the differences and similarities among different eukaryotes.

2. RNA interference as a technical tool

Until recently, the use of RNAi as a reverse genetic tool to study gene function was restricted to plants, Caenorhabditis elegans and Drosophila, where large dsRNAs can efficiently induce gene-specific silencing.

In mammalian cells, long dsRNA molecules (> 30 bp) lead to the activation of a dsRNA-dependent protein kinase (PKR) by dsRNA. PKR activation leads to a generalized inhibition of translation as well as to the induction of apoptosis. Exposure of cells to dsRNA can lead to activation of the type 1 interferon response and the STAT-mediated expression of PKR, hence amplifying the response.
Activated PKR leads to the phosphorylation of the small subunit of the eukaryotic initiation factor2 (eIF2) resulting in a global shutdown of translation. dsRNA also promotes the synthesis of 2'–5' polyadenylic acid, which, in turn, activates the non-specific Rnase L. These collective phenomena can dramatically alter cellular metabolism and often activate apoptotic pathways.
Therefore, transfection of large dsRNA is not the gene-silencing tool of choice in higher organisms. However, short siRNA molecules are capable of specifically silencing gene expression without induction of the unspecific interferon response pathway. Thus, siRNAs have become a new and powerful alternative to other genetic tools such as antisense oligonucleotides and ribozymes to analyze loss-of-function phenotypes.

Effective strategies to deliver siRNAs to target cells in cell culture include transduction by physical or chemical transfection. An alternative strategy uses the endogenous expression of siRNAs by various Pol III promoter expression cassettes that allow transcription of functional siRNAs or their precursors.