0273-2289/04/119/0001/$25.00

RNA Interference Silencing the Transcriptional Message

Aspects and Applications

Shantanu Karkare,*[,1] Saurabha Daniel,[1] and Deepak Bhatnagar[2]

[1]Apticraft Systems (P) Ltd. 142, Electronics Complex, Indore 452010, India, E-mail: Shantanu@premierebiosoft.com; and [2]School of Biochemistry, D.A University, Khandwa Road, Indore 452017, India

Received March 12, 2004; Revised July 10, 2004; Accepted April 15, 2004

Abstract

RNA interference (RNAi) is silencing of gene expression by double-stranded RNA (dsRNA) having complementary sequence to the target gene to be silenced. This phenomenon has transformed into a complete technology for functional genomic studies. Small interfering RNAs (siRNAs) are 21- to 23-nucleotide dsRNAs, in which the sense strand is the same as the target mRNA and the antisense strand is the complement of the target mRNA sequence. These are the effector molecules for inducing RNAi, leading to posttranscriptional gene silencing with RNA-induced silencing complex. Besides siRNA, which can be chemically synthesized, various other systems in the form of potential effector molecules for posttranscriptional gene silencing are available, such as short hairpin RNAs (shRNAs), long dsRNAs, short temporal RNAs, and micro RNAs (miRNAs). These effector molecules either are processed into siRNA such as in the case of shRNA or directly aid gene silencing as in the case of miRNA. RNAi for various unknown genes may facilitate to elucidate inherited genetic diseases and provide drug candidates for viral and oncogenic diseases. This can be achieved by targeting mRNA from oncogenic genes or mRNA for viral cellular receptor and viral structural proteins for RNAi. In this article, we evaluate various aspects and applications of RNAi technology and provide comprehensive information for the system currently available for inducing RNAi.

Index Entries: Small interfering RNA; short hairpin RNA; micro RNA; RNA-induced silencing complex; siRNA expression cassettes; siRNA expression vectors.

*Author to whom all correspondence and reprint requests should be addressed.

Introduction

RNA interference (RNAi), once proclaimed to be a bizarre phenomenon limited to a few plant species such as petunias *(1)* for posttranscriptional gene silencing, has now found an important niche in molecular biology research. The technology was earlier termed cosuppression or posttranscriptional gene silencing (PTGS) in transgenic plants and quelling in fungi. The phenomenon and its use as technology can be appreciated with the understanding of its evolution, mechanism, and applications.

RNAi can be defined as silencing of gene expression by double-stranded RNA (dsRNA) having complementary sequence to the target gene to be silenced *(2)*. Expression studies in *Caenorhabditis elegans* led to the first evidence for gene-silencing activity of dsRNA *(3)*. During functional studies on *par-1* gene in *C. elegans*, it was observed that besides antisense RNA, sense RNA control also disrupted *par-1* gene expression. The problem was later solved by Fire et al. *(4)*, who injected dsRNA, a mixture of both sense and antisense RNA strands, into *C. elegans*. Effective gene silencing was observed with dsRNA compared to either sense or antisense strand. Through these and other related studies, RNAi has become a major technological breakthrough in the field of research, opening new avenues in functional genomics studies in drosophila, mouse, human, and mammalian cell system *(5)* to target-specific drug design.

Mechanism of RNAi

RNAi as a natural phenomenon is likely to protect organisms from viruses and suppress the activity of transposons. Biochemical and genetic studies have prompted the current and widely accepted model for RNAi. This model can be systematically discussed in three steps: initiation, effector, and amplification.

Initiation

The initiation step involves digestion of dsRNA into 21- to 23-nucleotides (nt), small interfering RNAs (siRNAs). The dsRNA is introduced directly, via a transgene or a virus that is processed by Dicer enzyme into siRNA, also known as the guide RNA by an adenosine triphosphate (ATP)–dependent process (Fig. 1). Dicer is a member of the RNase III family of dsRNA-specific ribonucleases. Dicer has been identified for various domains that form the functional entity of the enzyme. It has an N-terminal RNA helicase domain; a Piwi, Argonaute, Zwille/Pinhead as specific domains; two RNase III domains; and a C-terminal dsRNA-binding motif.

Genomic studies of the mechanism have revealed the participation of *rde1* and *rde4* (RNAi deficient) *C. elegans* gene in the initiation step of RNAi. *rde1* gene is a member of a large family of genes and is homologous to the Neurospora *qde2* (quelling deficient) and Arabidopsis *AGO1* genes (Argonaute, previously identified to be involved in development). The detailed function of these genes is, however, not clear *(6,7)*.

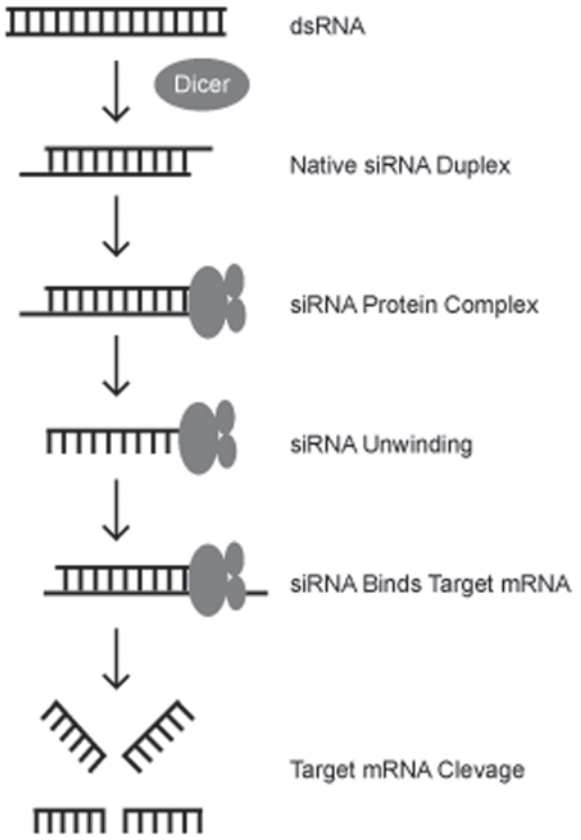

Fig. 1. RNAi mechanism. dsRNA is processed by Dicer RNase III into 21- to 23-nt siRNA duplexes. These siRNA duplexes are incorporated into a ribonucleoprotein complex to form the RISC endonuclease. This complex targets homologous mRNAs for degradation.

Effector

The effector step involves the formation of RNA-induced silencing complex (RISC) from siRNA duplex coherence to a nuclease complex. The active RISC targets the homologous transcript by base-pairing interaction. The activation of RISC is achieved by ATP-dependent unwinding of siRNA duplex (Fig. 1). *rde2* and *mut7* genes have been found to be important for the effector step in *C. elegans*. The identification of these effector genes was possible with studies on heterozygous mutant worms that were unable to transmit RNAi to their homozygous offspring *(6)*.

Amplification

The remarkable natural potency of RNAi in some organism has led to an amplification step within the RNAi pathway. siRNA should reach sufficient concentration to accomplish target mRNA cleavage. This is facilitated by target RNA copying steps in organisms such as *C. elegans* and

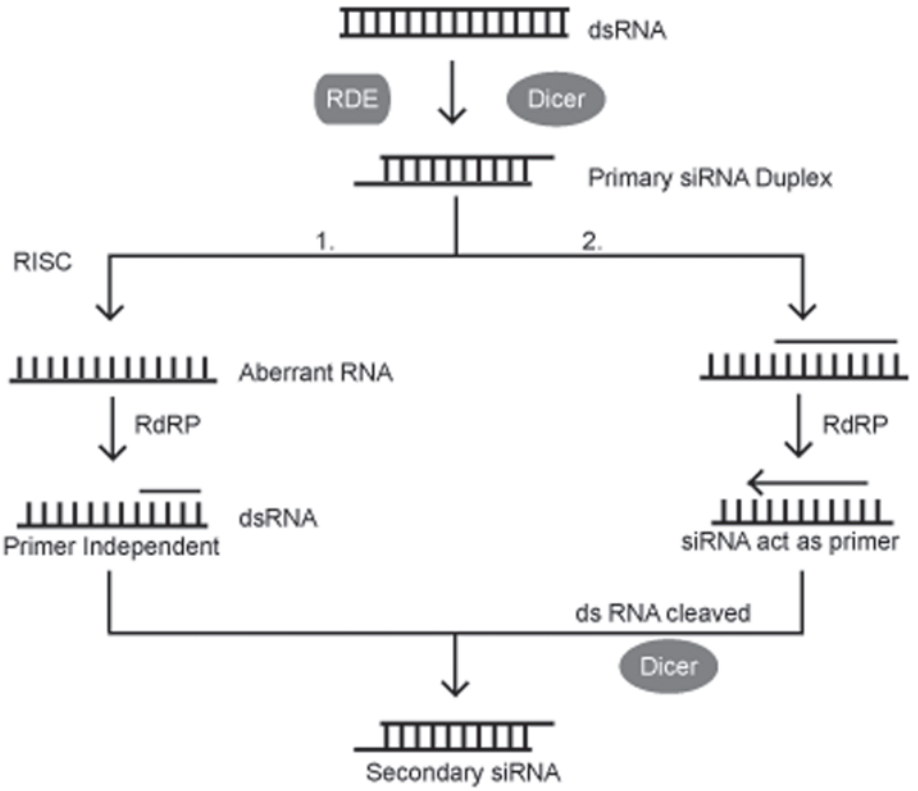

Fig. 2. Amplification in siRNA pathway. In an organism such as *C. elegans*, dsRNA is processed into siRNA with the help of Dicer and RDE. Then primary siRNA is (1) processed into aberrant RNA by RISC followed by primer-independent copying at the 3'end of aberrant RNA, resulting in siRNA, or (2) converted into dsRNA by RdRP by a primer-dependent process. The resulting siRNA is finally converted to secondary siRNA by Dicer.

Neurospora. *C. elegans* proves to be a good model for studying the amplification step because of the unstable nature of siRNA, which indicates the participation of RNA-dependent RNA polymerases (RdRPs) for generating secondary siRNAs, allowing sustained RNAi response. In *C. elegans*, two pathways are reported for siRNA in the amplification step: (1) a strand of siRNA is processed into aberrant RNA by RISC, which is followed by RdRP-initiated, primer-independent copying of the aberrant RNA into dsRNA; and (2) one of the siRNA strands as a primer along with RdRP generates dsRNA, escaping the RISC processing of siRNA. Finally, dsRNA resulting from RdRP copying of an aberrant transcript is processed by Dicer into secondary siRNAs (Fig. 2). These secondary siRNAs form a part of the RISC complex.

At the genomic level, Neurospora qde-1, Arabidopsis *SDE-1/SGS-2*, and *C. elegans ego-1*, appear to encode RdRP *(6)*.

siRNA and RNA Effecting RNAi

Advancement in the understanding of siRNA and concerns about the success of RNAi experiments have led to the generation of a range of systems from which researchers can choose. These systems can be viewed as various RNAi-based platforms, which directly employ siRNA or precursors of siRNA such as short hairpin (shRNA).

siRNA

Chemical synthesis of siRNA is one of the popular methods used for generating siRNA. It involves identification of potential target sites on the target sequences, which will be taken as templates for siRNA design. With the advent of Web-based tools for siRNA target identification and siRNA design (siRNA Target Finder and Design Tool), it is now possible to order on-line siRNAs for synthesis, once they have been designed. Much, emphasis is being placed on the selection of appropriate target sites *(8)* because this determines, to a large extent, the success of experiments. Among the various guidelines and improved proprietary algorithms available for siRNA design, the following guidelines provided by Tuschl and Borkhardt *(2)* form the basis of many improved algorithms:

1. The target regions are identified 50–100 nt downstream of the start codon in a given cDNA sequence. Generally, 5'- and 3'-untranslated regions near the start codon are avoided.
2. For designing siRNA, a duplex 23-nt motif (AA[N19]TT) is searched with GC-content in a range of 30–70%. It is suggested that if no suitable sequences are found, the search can be extended for the NA(N21) motif.
3. siRNA sequences of 21–23 nt in length have been found to be effective for gene silencing.
4. There are two strands in siRNA: sense siRNA and antisense siRNA. The sense siRNA sequence corresponds to (N19)TT. The TT at the 3' end of the sense strand facilitates a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhang, since TT is also a 3' overhang for antisense siRNA. TT is preferred as the 3' overhang to simplify chemical synthesis. The antisense siRNA is a complement to position 1–21 of the 23-nt motif. It is important to note that the penultimate nucleotide of the antisense siRNA should be complementary to the target sequence.
5. For ensuring the success of siRNA experiments, it is recommended that BLAST (NCBI database) search, the selected siRNA sequences, to ensure that only one specific gene is targeted.

shRNA

The shRNAs are fold-back stem loop structures that are processed into siRNA and, hence, act as a kind of precursor molecule for siRNA. These

shRNAs can be generated in cell lines with the help of commercially available shRNA expression vectors such as pSilencer 2.0-U6 and 3.0-H1 vectors from Ambion, psiRNA-hH1, psiRNA-hH1neo, and psiRNA-hH1zeo vectors from Invivogen (Fig. 3), and siRNA expression cassettes (SECs) from Ambion's Silencer™ Express system.

The efficient and long-term suppression of gene expression is achieved with shRNA expression vectors, which prove to be an advantage over synthetic siRNAs. Cells that are transfected with siRNA expression vector experience steady and long-term mRNA inhibition, whereas cells that are transfected with exogenous synthetic siRNA typically recover from mRNA suppression within 7 d or 10 rounds of cell division. This long-term gene-silencing ability of siRNA expression vectors provides an ample opportunity for its applications in gene therapy.

The shRNA expression vectors are engineered plasmid vectors containing promoters of the type III class of Pol III promoters (H1 RNA, U 6 promoter), a cloning site for stem-looped RNA insert, and a 4 or 5-thymidine transcription termination signal. The polymerase III promoters are preferred because of their well-defined initiation and stop sites, besides the advantage that the transcripts lack a poly (A) tail. Five thymidines define the termination signal for these promoters, and transcript is cleaved after the second uridine, which generates the 3' UU overhang in expressed siRNA, similar to the 3' overhang of synthetic siRNA.

Paddison et al. *(5)* studied the importance of stem and loop length, sequence specificity, and the presence of overhangs, important factors in determining shRNA activity, and came up with the following results:

1. Stem lengths can range anywhere from 25 to 29 nt and loop size between 4 and 23 nt without affecting silencing activity.
2. There was no decrease in potency owing to the presence of G-U mismatches between the two strands of the shRNA stem.
3. Complementarities between the portions of the stem that bind to the target mRNA are important for the success of the experiment. The presence of single-base mismatches between the strand of the stem and the mRNA abolished silencing.
4. The orientation of insertion into the pol III expression vectors is an important criterion.

Although variations are observed in the design consideration of shRNA insert, the basic configuration of the insert includes ~20-nt inverted repeats that code for dsRNA stem complementary to the target gene and ~10-nt spacers that code for the loop in hairpin. To design an shRNA insert, it is necessary to start with the identification of the target in the gene sequence. The same criteria applicable for the siRNA design can be employed for selecting the target in sequence followed by specific guidelines specified for the shRNA insert design by the manufacturers that provide various shRNA expression vectors. The following considerations are applicable as general guidelines for designing an shRNA insert:

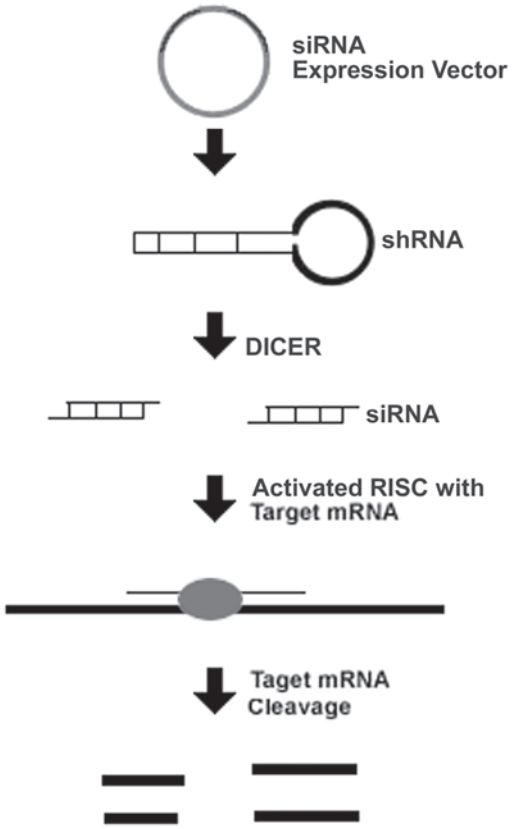

Fig. 3. shRNA vector–based siRNA technology. DNA insert (about 70 bp) that encodes for shRNA targeting the gene of interest is cloned into a vector. On vector expression, this shRNA is rapidly processed by a cellular mechanism in vivo into siRNA.

1. The insert should have vector-specific suitable restriction overhangs.
2. The length of target-specific duplex and loop must be considered for effective silencing.
3. The first base of the shRNA target-specific sequence should start with nucleotides that correspond to the native transcription start site for the particular promoter.
4. The repeats in the shRNA insert must be avoided because they may result in premature transcription termination of shRNA.

Besides the commercially available vectors from the various manufacturers, customized vectors can also be generated depending on the research requirements.

Another alternative for shRNA generation is SECs. SECs are polymerase chain reaction (PCR) products consisting of promoter and terminator sequences flanking a hairpin siRNA template. Gene-specific silencing is effected by the hairpin siRNA expressed from the PCR product. SECs

circumvent the following drawbacks of vector-based systems for shRNA generation:

1. SECs can be transfected into a broad range of mammalian cells with high efficiency. They are small, 160–360 bp, depending on the length of RNA polymerase promoters, whereas in vector-based systems, transfection efficiencies of plasmids tend to be relatively low owing to the size of the plasmid. This can make it difficult to use plasmids in homogeneous assays under nonselective conditions, especially with difficult-to-transfect cells.
2. SECs require two to three PCRs with target-specific oligonucleotides for creating products for transfection, which takes less time compared to the long and laborious process of preparing vectors.
3. To identify combinations of promoter and siRNA sequences that function best in the experimental system, multiple SECs can be generated in a short time for each mRNA target. The best SECs then can be employed for the vector system, whereas using shRNA expression vectors for screening, potent siRNAs that can induce RNAi, is expensive and time-consuming.

The process for generating an SEC can be outlined in three steps:

1. Precursor SEC generation involves one to two PCRs using promoter elements, promoter PCR primers, and one to two gene-specific oligonucleotides. Ambion's Silencer™ Express system allows precursor SEC generation from single-oligonucleotide template and from sense and antisense siRNA template oligonucleotide using either a single PCR or two consecutive PCRs.
2. SEC is generated from precursor SEC as template.
3. SEC can be purified by affinity columns to remove contaminating oligonucleotides, dNTPs, enzymes, and salts.

Long dsRNA

RNAi experiments in nonmammalian targets can be performed with dsRNA of 400 bp or larger. The long dsRNA was found to mediate gene silencing during initial studies on *C. elegans*. dsRNA is processed by a ribonuclease III–like enzyme into siRNAs. The minimum size of dsRNA recommended for RNAi is ~200 bp. It is suggested that longer dsRNA molecules are more effective on a molar basis at silencing protein expression, but a higher concentration of smaller dsRNA molecules may also have similar silencing effects.

The DNA templates for transcription of dsRNA can be generated by PCR. Various kit systems are available for designing these DNA templates, such as T7 RiboMAX™ Express RNAi System from Promega and Knockdown siRNA Kit from Spring Bioscience. The basic configuration of DNA template consists of promoter sequences at the 5' end of sense and antisense strands that will be translated into long dsRNA. The protocols for designing these templates varies with the kits and systems provided commercially.

Furthermore, expression cassettes can be employed for designing siRNA templates (siLentGene™ U6 Cassette RNA Interference System from Promega; Silencer™ siRNA from Ambion). There are two variant forms, one employing T7 promoter and the other employing U6 promoter cassette. siRNA targets (for both sense and antisense siRNA) will have leader sequence for T7 promoter at the 3' end (Silencer™ siRNA) or U6 cassette matching sequence at the 3' end and terminator sequence (siLentGene™ U6 Cassette RNA Interference System) so that these act as PCR primers that generate sense and antisense templates.

Small Temporal RNA

Small temporal RNAs are also regulatory small RNAs of 22 nt; lin-4 and let-7 RNAs are examples of this group. Initially processed from an ~70-nt single-stranded RNA transcript folded into a stem-loop structure, they prevent translation of their target mRNAs by binding to the target's complementary 3'-untranslated regions *(9)*.

Micro RNAs

Micro RNAs (miRNAs) are a novel family of small RNAs of 22 nt with potential roles in gene regulation. In terms of biochemical properties, they are the same as siRNA. miRNA sequences are found in the stem of a stem-loop structure *(7)*. In the nucleus, pri-miRNAs are generated that are trimmed into the characteristic pre-miRNA (70-nt stem-loop form). However, RNA polymerase and nuclear processing enzyme have not been identified. The pre-miRNA with the help of Dicer enzyme is processed into 22-nt mature miRNAs.

miRNA studies were initially carried out on lin-4 and let-7 RNA identified during genetic analysis of *C. elegans* developmental trimmings. It is suggested that miRNA mediates translation inhibition, and it is being investigated as a participant in a wide range of regulatory pathways *(10)*.

Applications of RNAi

The advent of RNAi technology has brought a revolution in molecular genetics and provided an efficient tool for the study of functional genomics. The vast applications and the utility of this technology are as follows:

Functional Genomics

RNAi and PTGS have proved to be potential tools for gene function studies. Biologists can switch off a specific gene in a variety of organisms in order to deduce gene functions with the help of PTGS. siRNA can be generated corresponding to the gene under study, followed by identification of phenotypic changes after transfection. These phenotypic changes can be in terms of viability or biochemical changes. This will provide a clue to the probable role played by the gene and will also provide a paradigm

shift in attempts to alter traits for improved agricultural productivity such as slowing the production of protein that causes fruits to ripen. The dsRNA or a DNA (expression cassettes or expression vectors) corresponding to the gene of interest is introduced into a cell. Either the shRNA is generated, which is processed into the amination siRNA, or the dsRNA activates the DICER/RISC process so that the properties of the affected cell reflect a loss of function in the corresponding gene.

Extensive research is being carried out to determine an efficient mode that can be employed to induce RNAi in various organisms. In the case of *C. elegans*, the nematode was fed with *Escherichia coli* strains that had been generated to produce dsRNA corresponding to the gene of interest, besides injection of dsRNA in the gonads or soaking in dsRNA *(6)*. A bacterial library for inactivation of 16,757 of a predicted 19,757 genes of worm was developed followed by phenotypic studies. Similarly, plants were infected with viruses that contained inserts corresponding to the genes of interest in the plant genome. Plant viruses are known to induce viral-induced gene silencing.

Drosophila, which does not exhibit systematic RNAi because it does not spread to other cells and tissues, provides an opportunity for researchers to perform cell-specific RNAi. In the case of cultured mammalian cells, RNAi is also becoming a standard laboratory technique that will be useful for human genomewide functional genomics.

Medical Applications: From Research to Therapy

Biomedical research aimed at exploring the potential of RNAi as an optimal therapeutic tool for various viral infections, cancers, and inherited genetic disorders has been conducted *(11,12)*.

Major viruses such as human immunodeficiency virus, human papilloma virus (HPV), and hepatitis B virus are the targets for RNAi-based therapeutics. Viral genes such as *gag* have been silenced for inhibiting viral replication in cultured cells besides cellular genes such as CD4 necessary for viral infection *(13)*. The single-stranded RNA of hepatitis C virus can be a potential target for RNAi. Similarly, the silencing of E6 and E7 genes has been successful in HPV type 1 *(14)*.

One of the outstanding features of siRNA that is desired in any good drug is the specific nature of siRNA for its target. This potential of siRNA can be exploited for silencing mutated genes. This has been tested with *K-RAS* oncogene where loss of tumorigenicity was observed on loss of expression. The use of a retroviral version of an H1 promoter-driven shRNA expression system has been reported to be successful for inhibition of expression of mutated *K-RAS* in both tissue culture and animal model *(15)*. *M-BCR/ABL* fusion, *c-raf* and *bd-2*, and vascular endothelial growth factor are also possible targets for tumor therapy by siRNA *(2)*.

RNAi also provides a therapeutic role in inherited diseases and human neurodegenerative disorders such as Huntington disease and Kennedy

disease. The elucidation of signal transduction pathways is also an application of siRNA targeted at the intermediates of cascade.

In the pharmaceutical industry, there is always a bulk of candidate drug targets for which drugs can be developed. This is making the selection of bona fide drug targets more critical. Failed drugs that block the target protein but do not alleviate the disease state account for the majority of the investment in drug development. Successful drug target validation in vitro can be made possible by knocking down target gene expression effectively and with high specificity.

Even with the ample potential in RNAi technology, a few hurdles still remain on its way to routine clinical use. One of the major concerns in this regard is the efficient delivery of a sufficient quantity of siRNA into a significant number of target cells. Associated with this is the problem of targeted administration of siRNA, difficult with synthetic siRNAs. Another limitation is sequence specificity. Frequent mutation in disease genes will nullify the effect of siRNA.

Plant Pathology and Agricultural Production

In plants, RNA silencing is used as a defense mechanism against viral infections. It can confer immunity against closely related viruses. There are two types of plant viruses containing genomic RNA or DNA. Those with RNA genomes are strong inducers of RNA silencing since a dsRNA is formed as an intermediate during replication. This technology can circumvent the problem of crop loss owing to viral infection, which will increase crop productivity.

Transgenic tomatoes resistant to Crown gall disease have been generated, which highlights the potency of this technology in improving agricultural production. These plants are engineered to contain versions of two bacterial genes causing the disease. The spread of infection is thus prevented when the extra gene recognizes and shuts down expression of the corresponding bacterial gene during infection *(6)*.

Acknowledgment

We wish to thank Mr. Arun Apte, Chief Technology Officer, PREMIER Biosoft International for his guidance and encouragement during this work.

References

1 Jorgensen, R. A., Cluster, P. D., English, J., Que, Q., and Napoli, C. A. (1996), *Plant Mol. Biol.* **31,** 957–973.
2 Tuschl, T. and Borkhardt, A. (2002), *Mol. Interventions* **2,** 158–167.
3 Guo, S. and Kempheus, K. J. (1995), *Cell* **81,** 611–620.
4 Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998), *Nature* **391,** 806–811.
5 Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., and Conklin, D. S. (2002), *Genes Dev.* **16,** 948–958.

6. Thakur, A. (2003), *Electron. J. Biotechnol.* **6,** 39–49.
7 Tuschl, T. (2001), *Chem. Biochem.* **2,** 239–245.
8 Jarvis, R. A. and Ford, L. P. (2001), *Ambion Tech. Notes* **8,** 3–5.
9 Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001), *Cell* **106,** 23–34.
10 Grosshans, H. and Slack, F. J. (2002), *J. Cell Biol.* **156,** 17–21.
11 Brummelkamp, T. R., Bernards, R., and Agami, R. (2002), *Science* **296,** 550–552.
12 Tuschl, T. (2002), *Nat. Biotechnol.* **20,** 446–448.
13. Carl, D. N., Michael, F. M., Derek, M. D., and Paul, J. (2002), *Nat. Med.* **8,** 681–686.
14 Kim, N. V. (2003), *J. Korean Med. Sci.* **18,** 309–318.
15. Brummelkamp, T. R., Bernards, R., and Agami, R. (2002), *Cancer Cell* **2(3),** 243–247.