1

Principles of Dicer Substrate (D-siRNA) Design and Function

Mohammed Amarzguioui and John J. Rossi

Summary

An efficient RNAi largely depends on optimal design of the siRNA. In recent studies, Dicer substrates were found to be more potent than classical synthetic 21-mer siRNAs, suggesting a coupling of the Dicer-mediated processing step to the efficient assembly of the silencing complex, RISC. We describe the fundamental principles and experimental results leading to optimal Dicer substrates.

Key Words: siRNA; shRNA; RNAi; Dicer.

1. Introduction

The formal description of RNAi as a biological response to double-stranded RNA resulted from a desire to understand a number of intriguing observations arising from the use of antisense RNAs in *C. elegans* ***(1)***. This ultimately led to the discovery that worms could be programmed to silence genes by exposing animals to homologous dsRNAs (termed *triggers*) ***(1)***. It is now clear that an RNAi pathway is present in many if not most eukaryotes ***(2)***. A biochemical understanding of the RNAi pathway was crucial to realizing that dsRNAs shorter than 30 bp could be used to specifically trigger an RNAi response in mammals. Tuschl and colleagues demonstrated that transfection of mammalian cells with small interfering RNAs (siRNAs) could specifically induce RNAi and thus cracked the barrier to the use of RNAi as a genetic tool in mammals ***(3)***. It took a remarkably short period of time for siRNAs to be adopted as a standard component of the molecular biology toolkit.

From: *Methods in Molecular Biology, vol. 442: RNAi: Design and Application*
Edited by: S. Barik © Humana Press, Totowa, NJ

The introduction of siRNAs into mammalian cells can be achieved through standard transfection. The strength and duration of the silencing response are determined by several factors. On a population basis, the overall efficiency of transfection is a major determinant, which must be addressed by optimizing conditions. In each individual cell, silencing depends upon a combination of the amount of siRNA that is delivered and upon the potential of that individual siRNA to suppress its target (the potency). Even a relatively poor siRNA can silence its target provided that sufficient quantities are delivered. However, essentially "forcing" the system with such reagents is likely to lead to numerous undesired effects ***(4)***.

Most investigators utilize siRNAs that consist of 19-base duplexes and 2-base 3′-overhangs on each strand. Our laboratory and that of Greg Hannon have both reported improved efficacy of longer-than-standard RNAi effectors. During investigation of cellular interferon induction caused by *in vitro* transcribed siRNAs, it was observed that at limiting concentrations some siRNAs of length 25–27 appeared to have greater potency than 21-mer siRNAs that could potentially be generated from the larger duplex ***(5)***. Hannon and colleagues reported a similar phenomenon for shRNA ***(6)***. They found synthetic shRNAs with 29-base-pair stems and 2-nucleotide 3′ overhangs to be more potent inducers of RNAi than shorter hairpins ***(6)***. Their studies further demonstrated that *in vitro* processing by Dicer was directional, starting predominantly from the open end of the stem and generating a mixture of 21- and 22-mer cleavage products. In both above cases, the increased potency can confidently be attributed to Dicer processing, which is thought to promote more efficient incorporation into RISC through the physical association of Dicer with the effectors of RNAi, the Argonaute proteins. This interpretation is supported by biochemical evidence in the fruit fly indicating a role for Dicer in the initial stages of RISC assembly ***(7)*** and by recent reports that Dicer-mediated processing of miRNA precursors in human cells is functionally coupled to miRISC assembly and improves subsequent silencing ***(8,9)***.

2. Materials

1. An algorithm for the prediction of relevant parameters of siRNA such as secondary structure, thermodynamic property, etc. (*See* the algorithm provided in **Section 3**).
2. In-house RNA synthetic facility or a commercial source (such as Dharmacon, Ambion, IDT, Sigma-Proligo, etc.) for obtaining siRNA of various lengths and compositions.
3. An assay for siRNA function, primarily a cultured cell line with a target gene.
4. If stable knockdown is desired, express the Dicer substrate from recombinant lentiviral or other suitable vectors (*see* **Chapter 8**, for example).

3. Methods

3.1. Design of D-siRNAs

While Dicer processing is generally beneficial, the composition and potency of the processing products are also of importance for the overall efficacy. Dicer processing of unmodified 27-mer duplexes is largely unpredictable, sometimes resulting in the generation of siRNAs of poor activity, thereby reducing the activity of the 27-mer to below that of an optimal 21-mer within its target sequence. Consequently, there is no guarantee that a randomly designed 27-mer will be more efficacious than the best of the potential 21-mers within its target sequence. The problem of making the processing of Dicer substrate siRNA (D-siRNA) predictable, thereby enabling rational design on the basis of published design algorithms, now appears to have been solved. The new optimal design introduces directionality and uniqueness of processing into the Dicer cleavage step by mimicking the relevant structural features of naturally occurring Dicer substrates, pre-miRNAs. These can be summarized as follows.

1. To select the D-siRNA sequence, identify potentially good siRNA/target combinations using a computer-based algorithm. We have one that analyzes all the possible 21-mer targets in a given mRNA for secondary structure, GC content, and thermodynamic end properties of the 21-mer siRNAs that would be used to target the sequence. The target/siRNAs are ranked top to bottom by a numerical value ***(13)***. Access to this algorithm is free using the following Web address: http://www.cityofhope.org/researchers/RossiJohn/RossiJohnResearch.htm. Provide an E-mail address to which the output, which can be accessed by Microsoft Excel, will be returned.
2. Have bulged stem-loop structures with 2-nt 3′-overhangs. Recent reports ***(6,10)*** suggest that the overhangs in the open end of the stem in such structures are bound by Dicer and determine the direction of processing as well as preferential strand selection.
3. When synthesizing a linear dsRNA substrate, make one end of the duplex blunt-ended to mimic a natural Dicer substrate in which the corresponding end of the duplex is closed by a loop, precluding binding of Dicer to that end. In addition, this feature introduces two DNA nucleotides in the sense strand in the blunt end of the duplex **(Fig. 1)** ***(11)***. The incorporation of a 3′-overhang in one end introduces a preference for processing to start from that end, while DNA nucleotides in the opposite blunt end enforce this asymmetry while blocking processing events involving the terminal two phosphodiester linkages. This results in the predictable production of a single or major 21-mer product of processing starting from the overhang terminus, sometimes accompanied by a minor 22-mer product resulting from processing from the same end. This mixture is similar to that previously reported to result from Dicer cleavage ***(6,12)***, and this flexibility in Dicer processing may reflect some level of sequence preference near the putative cleavage site.

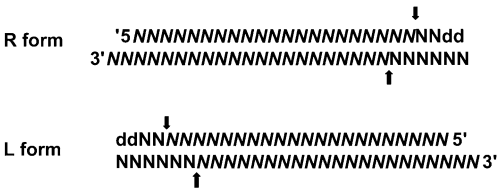
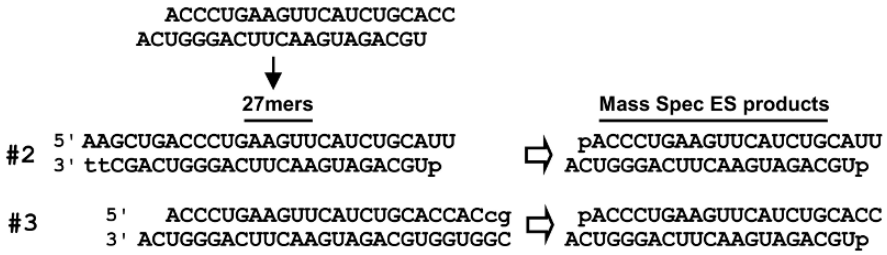


Fig. 1. D-siRNA conformations. Dicer entry is from the 2-base 3′-overhang. N = ribonucleic acid; d = deoxyribonucleic acid. Arrows point to Dicer cleavage sites, and italicized bases are the 21-mer siRNAs produced.

4. The same 21-mer can be generated from Dicer substrate siRNAs of slightly different sequence and opposite "polarity": one in which the passenger strand carries the 3′ overhang and processing proceeds from right to left (L-form), and the other in which the overhang is on the guide strand, and processing proceeds from left to right (R-form) **(Fig. 1)**. Interestingly, although processing of the two forms of D-siRNA produces the same 21-mer siRNA species (confirmed by mass spectrometry data; *see* **Fig. 2)**, when considering that the strand with the 2-base 3′-overhang is antisense to the target mRNA, the "R"-versions are consistently more efficacious than the "L"-forms. This comparison was performed for nine different pairs of D-siRNA, targeting four different genes. In seven out of nine cases, the R-form was superior to the L-form ***(11)***. We hypothesize that preferential binding to the 3′-overhang by Dicer during processing favors incorporation of the strand bearing the overhang. Thus, sense target silencing is more efficient with the R-form because that's the configuration in which the guide strand bears the overhang. This effect can be demonstrated experimentally by cotransfection experiments with reporters in which the two forms of D-siRNA (R and L) are cotransfected with a target gene **(Fig. 2)**. Silencing of the reporter is markedly better when the R-form is used even though the Mass Spec-determined siRNAs are virtually the same. Thus, the polarity of the Dicer entry is an important determinant for the selection of the guide strand, which is generally the strand with the 2-base 3′-overhang. This polarity also has the advantage of lower off-target possibilities by the passenger strand incorporation into RISC.
5. Due to variable efficacy of even rationally designed siRNAs, it is advisable to design multiple D-siRNAs targeting different sites and to titer their concentration to determine the optimal sequences and concentration for adequate silencing. Irrelevant control siRNA(s) should be included at all concentrations tested.
6. Avoid nonspecific and off-target effects (*see* **Notes 1 and 2**): Published data suggest that although near-complete inactivation of siRNA by a single mutation is possible, multiple mutations are generally necessary to ensure that the siRNA will be inactive ***(14–18)***. When targeting the 3′ UTR of a transcript, it is also important to consider the possibility that mismatches with the target may abrogate cleavage but can still function in translational repression. Although sequence context, mismatch type, and mismatch position all influence the impact of mutations



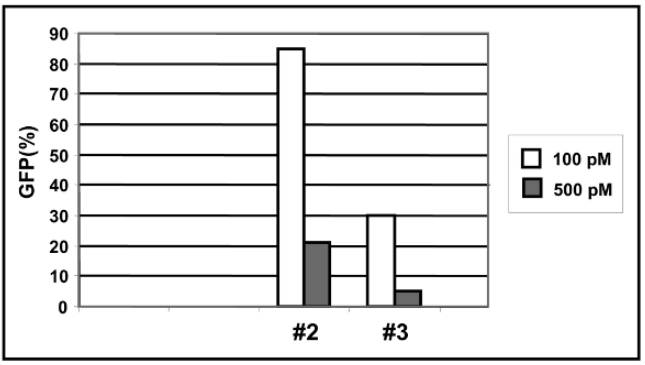


Fig. 2. Polarity of D-siRNA dictates efficacy. An EGFP target was cotransfected into HEK293 cells with either an L-form D-siRNA (#2) or an R-form (#3). The deoxynucleotides ("tt" in #2 and "cg" in #3) are in lowercase. The knockdown efficacy at two different concentrations of D-siRNAs was determined. Note that the R-form is the most potent for this target despite the fact that Dicer produces an identical bottom strand in both cases. The polarity of the Dicer entry is therefore an important determinant for guide strand selection into RISC.

(14,15,19), making it difficult to devise clear rules, some general guidelines can be formulated:

(a) Ideally, the selected siRNA should have multiple mismatches to all nontarget mRNA sequences.
(b) Mismatches located near the cleavage site or within the seed region (positions 2–11 within the putative guide strand and 9–18 within the target) are more disruptive than mismatches within the 5′ end of the target site. Full match within positions 9–18 of the "off-targets" should be avoided even if there are multiple mismatches in the 5′ end.
(c) Extensive matches to the self-complement of the target sequence are less critical since they would be targeted by the passenger strand, whose incorporation into RISC is designed to be minimal. For these sequences, the duplex region mismatch sensitivity is reversed (mismatches within positions 2–11 of the query are more critical).

7. Delivery of D-siRNA: Formation of lipid-siRNA or D-siRNA complexes should be performed in one batch for all different treatments of the same siRNA, and the complexes diluted appropriately to their final concentration ***(17)***. Adherent cells can be transfected either while adherent or in suspension after trypsin-mediated detachment ***(26)***. The latter procedure is recommended due to its greater flexibility and robustness (confluency of cells is not an issue). For difficult-to-transfect adherent cells, the methodology is frequently also associated with improved silencing ***(26)***.
8. Summary (*see* **Note 3**): Based on the aforesaid, once the optimal 21-mers are chosen using an algorithm, extend the passenger strand to 25 bases and make the last two nucleotides deoxys. Similarly, extend the guide strand to 27 nucleotides with (**Fig. 1**). Deliver the final D-siRNA using cationic lipids or electroporation and assay for loss of target. Screen various amounts of each D-siRNA for the most efficient knockdown.

4. Notes

1. Off-target effects: siRNA specificity determinations have traditionally been performed using BLAST searches. Recent data, however, have cast serious doubt on the value of BLAST searches for general siRNA specificity determination ***(19)***. Experimentally determined siRNA off-target effects were shown to correlate strongly with matches between positions 2–8 within the guide strand (the seed sequence) and sequences in the 3′ UTRs of affected genes ***(19)***. These seed matches are too short to be confidently detected by BLAST. These searches are therefore only useful for identifying near-perfect matches. A Web-based search tool is available for identification of all possible seed matches for any given siRNA (http://www.dharmacon.com/seedlocator/default.aspx). Despite the significant correlation of seed matches with off-targets, the predictive value of such a list is at present limited since only a small fraction of seed matches results in actual off-target effects. Thus, while BLAST searches can be used to select away the poorest candidates, and seed match searches can help in deciding among sequences on the basis of the number of potential off-targets, there appears for the foreseeable future no substitute for experimental determination of specificity, preferably by genome-wide gene expression profiling. Thus, for functional genomic studies, verification of phenotypes by the application of a combination of multiple active, target-specific siRNAs as well as inactivated or irrelevant control sequences is of paramount importance. Since D-siRNAs are processed to predictable 21- and 22-mer sequences, the off-target effects are not expected to be substantially different from those resulting from their corresponding 21-mers.
2. Activation of innate immune response: An emerging property of siRNA is their potentially immune-stimulatory effects *in vitro* and *in vivo* as a result of engaging members of the Toll-like receptor family following liposome-mediated endosomal trafficking ***(20–24)***. No induction of interferon or activation of PKR was observed in HEK293 cells following delivery of 27-mer siRNA ***(5)***. It should be pointed

out that in some cells the D-siRNAs have been shown to activate the antiviral protein RIG1. This activation appears to be eliminated ***(25)*** by the design shown in **Figs. 1** and **2**.

3. The design and application of D-siRNAs do not differ significantly from the standard 21-mer siRNAs. The advantages of the D-siRNA design described above are better selectivity of the guide strand as a consequence of Dicer processing and handoff to RISC and, in many instances, improved potency, which translates into lower effective concentrations. Other potential advantages of D-siRNAs are that they can be physically linked to cell-specific delivery ligands such as aptamers ***(27)*** or other nonnucleic acid compounds, and the siRNAs can be processed from the longer molecules via interaction of Dicer with the 2-base 3′-overhang in the D-siRNA design.

References

1. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391,** 806–811.
2. Hannon, G. J., and Rossi, J. J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature* **431,** 371–378.
3. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 494–498.
4. Jackson, A. L., Bartz, S. R., Schelter, J., et al. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol*. **21,** 635–637.
5. Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S., and Rossi, J. J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol*. **23,** 222–226.
6. Siolas, D., Lerner, C., Burchard, J., et al. (2005). Synthetic shRNAs as potent RNAi triggers. *Nat. Biotechnol*. **23,** 227–231.
7. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W., and Sontheimer, E. J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* **117,** 83–94.
8. Gregory, R. I., Chendrimada, T. P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123,** 631–640.
9. Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev*. **19,** 2979–2990.
10. Krol, J., Sobczak, K., Wilczynska, U., et al. (2004). Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J. Biol. Chem*. **279,** 42230–42239.
11. Rose, S. D., Kim, D. H., Amarzguioui, M., et al. (2005). Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Res*. **33,** 4140–4156.

12. Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev*. **15,** 188–200.
13. Heale, B. S., Soifer, H. S., Bowers, C., and Rossi, J. J. (2005). siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Res*. **33,** e30.
14. Amarzguioui, M., Holen, T., Babaie, E., and Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res*. **31,** 589–595.
15. Du, Q., Thonberg, H., Wang, J., Wahlestedt, C., and Liang, Z. (2005). A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res*. **33,** 1671–1677.
16. Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J*. **20,** 6877–6888.
17. Holen, T., Amarzguioui, M., Wiiger, M. T., Babaie, E., and Prydz, H. (2002). Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res*. **30,** 1757–1766.
18. Miller, V. M., Gouvion, C. M., Davidson, B. L., and Paulson, H. L. (2004). Targeting Alzheimer's disease genes with RNA interference: An efficient strategy for silencing mutant alleles. *Nucleic Acids Res*. **32,** 661–668.
19. Birmingham, A., Anderson, E. M., Reynolds, A., et al. (2006). 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Meth*. **3,** 199–204.
20. Hornung, V., Guenthner-Biller, M., Bourquin, C., et al. (2005). Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat. Med*. **11,** 263–270.
21. Judge, A. D., Sood, V., Shaw, J. R., Fang, D., McClintock, K., and MacLachlan, I. (2005). Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat. Biotechnol*. **23,** 457–562.
22. Kariko, K., Bhuyan, P., Capodici, J., and Weissman, D. (2004). Small interfering RNAsmediate sequence-independent gene suppression and induce immune activation by signaling through Toll-like receptor 3. *J. Immunol*. **172,** 6545–6549.
23. Robbins, M. A., and Rossi, J. J. (2005). Sensing the danger in RNA. *Nat. Med*. **11,** 250–251.
24. Sioud, M. (2005). Induction of inflammatory cytokines and interferon responses by double- stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. *J. Mol. Biol*. **348,** 1079–1090.
25. Marques, J. T., Devosse, T., Wang, D., et al. (2006). A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nat. Biotechnol*. **24,** 559–565.
26. Amarzguioui, M. (2004). Improved siRNA-mediated silencing in refractory adherent cell lines by detachment and transfection in suspension. *Biotechniques* **36,** 766–768, 770.
27. Rossi, J. J. (2006). Partnering aptamer and RNAi technologies. *Mol. Ther.* **14,** 461–462.