
siRNA: Enhanced Functionality Through Rational Design and Chemical Modification

The authors give a comprehensive overview of the development of siRNA technology and discuss current research in the use of siRNA in gene knockout experiments and in the design of siRNA-based therapeutics. The article includes a comparison of synthetic, plasmid and viral delivery systems; a discussion about setting standard definitions of functionality and a look at the application of rational drug design principles to functional siRNA.

Jon Karpilow, Devin Leake and Bill Marshall

Discovery biology has been accelerated by the punctuated appearance of new and pioneering technologies. Innovations in high-speed cell sorting (1), polymerase chain reaction (PCR) (2), microscopy (3) and proteomics (4) all have been key to the dissection of basic biological questions and instrumental in the ongoing transformation of the medical diagnostic and therapeutic arenas. The recent identification of RNA interference (RNAi) marks the discovery of a ubiquitous and highly conserved biological process, the emergence of which could overshadow all related technological predecessors in terms of its potential applications to chromosomal mapping, gene function analysis, drug target validation and therapeutics. Unlike antisense and ribozyme technologies, RNAi quickly has moved from theoretical, chalkboard discussions to a reliable gene-knockdown tool. The versatility and relatively inexpensive nature of the procedure makes it accessible to laboratories with both specific and broad-based goals, thus leading many to predict that RNAi could soon replace more conventional procedures (e.g., the knockout mouse) as the primary method of inducing gene knockouts.

Merging Technologies

The emergence of RNAi as a research tool rivals the explosion of PCR in the 1980s and can be attributed to contributions from a minimum of three disparate fields. Foremost among these are the discoveries that illuminated RNAi as a key mechanism underlying post-transcriptional gene silencing. Modern references describing the phenotype in plant and fungal organisms date back

to just more than a decade ago (5). Working in petunias, Napoli et al., observed that attempts to enhance petal coloration by over-expression of a chalcone synthase (CHS) transgene led to an unexpected block in anthocyanin biosynthesis and the generation of a white petal phenotype of variable stability. Subsequent quantitation of CHS transcripts in white (affected) and violet (revertant) flowers showed that the endogenous levels of CHS mRNA were reduced by as much as 50-fold in affected tissues, and recovered to near normal levels in revertant plants, thus alerting researchers to the presence of a transient mechanism analogous to paramutation in maize (6). Two years later, Romano and Macino noted similar phenotypes (termed “quelling”) in *Neurospora crassa* isolates transformed with albino-3 and albino-1 gene fragments (7). As was the case in the petunia, quelling was found to be “spontaneously and progressively reversible,” again attesting to the transitory nature of the phenotype.

Craig Mello and Andrew Fire further elucidated this form of gene regulation while working with *Caenorhabditis elegans* (8). This group coined the term “RNA interference (RNAi)” when it observed that injection of double-stranded RNA sequences induced potent and sequence-specific gene silencing. The researchers documented that only a few molecules were required to extinguish the target message in each cell, thus suggesting that the process was catalytic in nature and thereby distinguishable from other forms of post-transcriptional gene silencing (e.g., antisense). Subsequent work by other groups — including those led by Tuschl, Zamore, Hannon and Plasterk — have contributed greatly toward understanding the biochemical nature of the



RNAi pathway (9–11). Specifically, current models hold that cellular introduction of long, double-stranded RNA (dsRNA) molecules leads to the recruitment of a Type-III-like RNase (Dicer) that cleaves the duplex into small (19–25 base-pairs) inhibitory duplexes (short interfering RNA, or siRNA). Subsequent incorporation of these molecules into a multisubunit — RNA-induced silencing complex (RISC) — is followed by recognition and destruction of complementary mRNAs via site-specific cleavage in the region of siRNA–mRNA homology.

In mammalian tissues, the presence of long dsRNA leads to an interferon response that overshadows the operations of RNAi. The dual activation of both pathways limits the utility of dsRNA as a trigger for post-transcriptional gene silencing in mammalian systems and alludes to a potential role of RNAi in cellular immunity (12). Fortunately, work by Elbashir et al., (13) demonstrated that the two functions could be segregated, with RNAi being activated by the introduction of synthetic siRNA molecules that successfully bypass the interferon response and enter the RNAi pathway downstream of Dicer-mediated cleavage.

While groundbreaking advancements in the field of post-transcriptional gene silencing clearly have been the lynchpin in the expansion of RNAi as a laboratory tool, improvements in nucleic acid chemistry also have been influential. The functional intermediate in RNAi is a small RNA duplex readily produced by a variety of enzymatic (*in vitro* transcription, digestion or *in situ* expression) and chemical synthesis strategies. Among these methods, commercially available chemical synthesis platforms offer the greatest reliability and flexibility. Traditional methods of synthesizing RNA were developed as an adaptation to DNA synthesis (e.g., 2'-O-t-butyldimethylsilyl, tBDMS; [14, 15], and 2'-O-triisopropylsilyloxymethyl, 2'TOM; [16]) and are sufficient for small-scale production of shorter oligonucleotides. However, these platforms are limited severely from the perspective of final yields, oligonucleotide purity and process scalability. The most recent improvement in the field of RNA synthesis chemistries is 2'-O-acetoxyethoxy (ACE®)-based chemistry, which addresses many of these shortcomings and represents the industry standard for synthetic RNA needs. Devised by Scaringe (17), 2'-O-ACE was the first chemistry developed specifically for RNA synthesis (Figure 1). The platform yields a water-soluble oligonucleotide that incorporates a nuclease-resistant, acid-labile, orthoester blocking group at the 2'-ribose position. The 2'-protecting group can be left intact after standard deprotection procedures and is removed easily using mildly acidic conditions. This provides a stabilized RNA intermediate with favorable handling properties that can be converted readily to standard RNA at a terminal step. More importantly, the 2'-O-ACE RNA chemistry is amenable to a wide range of modifications and exhibits coupling efficiencies that permit high yield synthesis of oligos of unprecedented lengths (80 bases or more). These features are key to the development of therapeutic molecules with desired serum stability, delivery and longevity traits and are essential for those interested in creating synthetic oligonucleotides without concern for length limitations.

Widespread availability of sequence data and bioinformatics tools also have played a pivotal role in the development of RNAi as a research and drug development tool. Genomic sequence and expression databases provide the basic foundation for creating

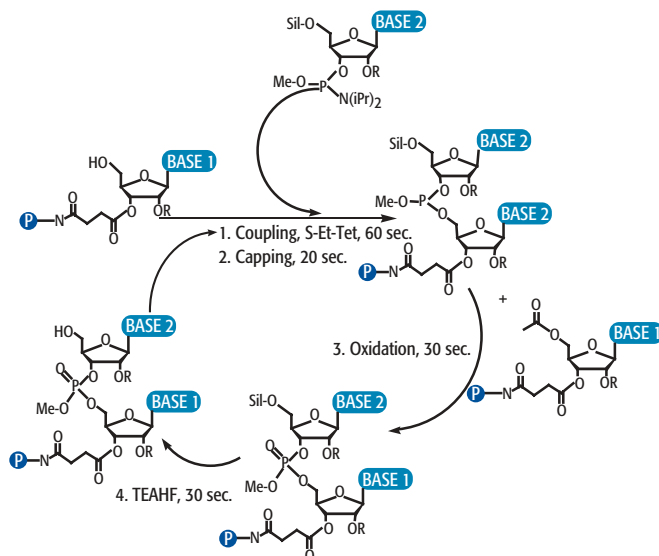


Figure 1. The RNA nucleotide synthesis addition cycle. For 2'-O-ACE chemical synthesis, the 5' position is protected by a silyl group and the 2' position is protected by an acetoxy ethyl orthoester group (OR).

siRNA design algorithms that enable researchers to efficiently and reliably identify regions of a gene (*in silico*) for inducing gene silencing. Equally important, high quality sequence data has played a vital role in ensuring siRNA specificity. Current, well-designed siRNA selection schemes include thorough BLAST™ (National Center for Biotechnology Information [NCBI], Bethesda, Maryland, USA) searches to minimize off-target effects generated by overlapping sequence homologies between the intended target and alternate splice forms, closely related family members (homologs) and unrelated genes that contain limited but conserved domains. Such bioinformatic procedures also limit off-target effects that result from sense strand homology with unintended targets. A comparison between two recently published articles shows that inclusion of a stringent BLAST search during siRNA design eliminated nearly all sense strand off-target effects (18, 19). In the absence of such procedures, numerous targets having as little as 11 nucleotides of homology to the sense strand of the siRNA, are down-regulated (18).

siRNA Design

In lower eukaryotes, target site selection is unnecessary given that Dicer processing of long dsRNA generates a heterogeneous population of siRNAs, one or more of which is capable of efficient gene silencing. The interferon response-based restrictions imposed upon higher eukaryotes prevent the exploitation of long dsRNA as an RNAi trigger in these organisms and places a greater responsibility on individual (or small pools of) siRNA duplexes. As siRNA-induced gene silencing appears to be independent of target secondary (or tertiary) structure, one fundamental challenge for successful implementation of RNAi rests on the ability to design duplexes that are both specific and potent.

Early studies that selected siRNA randomly or based upon a limited set of criteria yielded silencing efficiencies that were unpredictable. From today's perspective, this variability in silencing efficiencies is not surprising. The current model describing the RNAi

pathway involves no fewer than five distinct steps, including siRNA–RISC binding, duplex unwinding and strand selection, target identification, target cleavage and target release (Figure 2). As each step likely involves multiple protein–nucleic acid interactions, key sequence and biophysical parameters are expected to be essential at each stage to achieve optimum functionality. Obtaining quality siRNA designs is complicated further by the sheer amount of information that needs to be generated and processed during the design procedure. Encoded within a single three-kilobase (kb) transcript are 2981 potential siRNA targets (assuming 19 nucleotides [nt] per siRNA). Developing efficient bioinformatics tools that are capable of sifting through this set to identify functional duplexes is crucial to minimizing the time and costs associated with developing and screening research and therapeutic reagents.

One approach to identifying functional siRNAs combines data generated by *in silico* and empirical approaches. Using predefined, model targets (human cyclophilin B, human diazepam binding inhibitor [DBI] and firefly luciferase), synthetic siRNAs were prepared to every other position of specific regions of these genes and tested in cultured cells to identify functional and non-functional molecules (20). Subsequent analysis of these empirical results, with the particular intention of identifying positive and negative determinants of function, led to the identification of approximately 30 parameters that correlate with duplex functionality. In addition to refining optimal biophysical properties such as GC content (36–52%) and duplex thermal stability parameters, multiple sequence preferences at specific positions within the duplex were identified. Individually, these factors provided little advantage to a molecule's silencing ability; yet, when all of these factors were combined in a single, weighted algorithm, it was possible to analyze populations of sequences and segregate functional from non-functional targets with a very high probability (Figure 3). While this approach yielded a clear improvement over previous design strategies, it is important to note that the algorithm does not identify all functional duplexes, suggesting that additional key parameters for functionality still need to be identified.

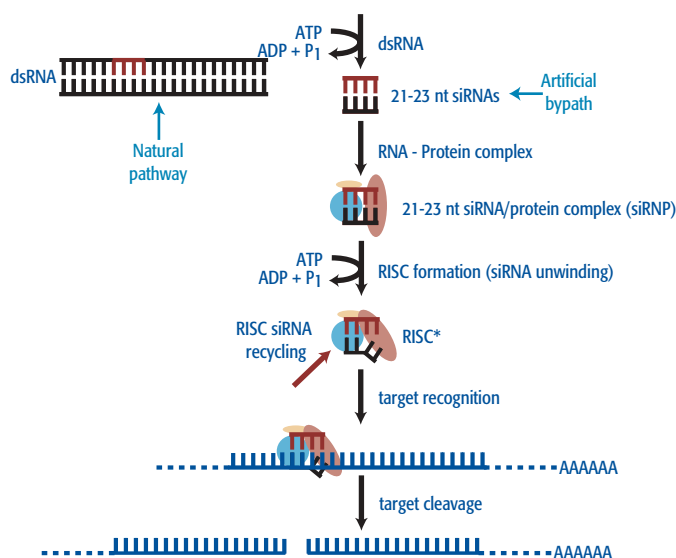


Figure 2. A mechanistic description of the steps involved in RNAi.

In addition to the aforementioned factors, analysis of the average internal stability profile of functional and non-functional duplexes identified a free energy profile that was characteristic of functional duplexes. Highly functional molecules exhibit a distinct, sequence-dependent, sinusoidal energy pattern (Figure 4). In contrast, the average stability profile of siRNA with poor silencing abilities was observed to be roughly 180° out of phase with its highly functional counterpart, thus suggesting that another dynamic parameter that contributes to functionality focuses on multiple regional stabilities within each duplex (21, 22).

What is the practical significance of the functional average internal stability profile? One prediction is that siRNAs exhibiting an optimal profile will interact preferentially with RISC and thus be protected (and preserved) in cells that otherwise are rife with nuclease activity. Interestingly, studies of siRNAs that are isolated from tissues transfected with long dsRNA show that only those duplexes exhibiting a functional average internal stability profile are retained, thus supporting the hypothesis that this thermodynamic profile contributes to duplex preservation (23).

The Concept of Pooling

Early reports in the siRNA literature hinted that introduction of multiple, randomly selected siRNA sequences directed against a given target dampened the observed level of silencing, possibly through a mechanism of competition between functional and non-functional sequences for a limited reservoir of RISC (24–27). These findings conflict with several preconceived notions of how the RNAi pathway is believed to function in nature. Foremost among these is the knowledge that cytoplasmic introduction of long, double-stranded RNA sequences — either by viral infection or transposon expansion — is expected to generate a diverse group of sequences, many of which are non-functional. As organisms that do not exhibit the interferon response are capable of dealing with such molecular assaults without succumbing to the effects of infection, it is reasonable to assume that one or more mechanisms are in place to deal with

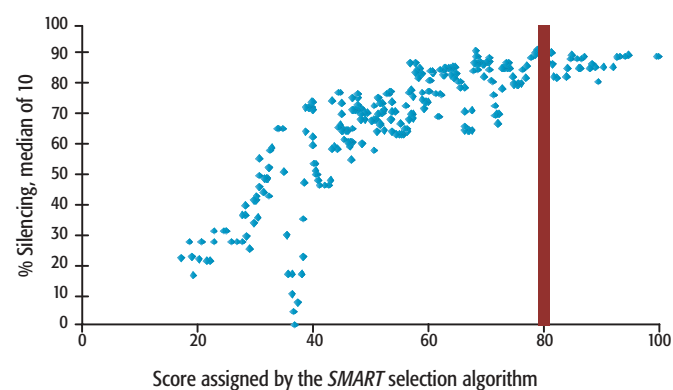


Figure 3. A test of 270 siRNAs directed against the human cyclophilin B, human DBI and firefly luciferase genes was analyzed using Dharmacon's rational design algorithm (SMART Selection technology). Subsequently, each duplex was transfected into cells and tested for the ability to silence the intended target. The plot shows the relationship between functionality and scores assigned to each siRNA by the algorithm.

the predicted competition between functional and non-functional duplexes. Previously discussed work in plants provides one explanation for how competition is avoided (21, 28): nature preserves only those siRNAs that have biophysical characteristics similar to functional siRNAs identified by rational design algorithms and discards those sequences that exhibit “non-functional” profiles. Thus, in instances where populations of functional and non-functional siRNA are present at non-saturating concentrations, nature eliminates competitive, non-functional sequences that diminish silencing. In contrast, under standard transfection conditions (i.e., lipid-mediated delivery, 100 nM siRNA) where RISC is saturated, the system might be incapable of removing the surplus of non-functional sequences, thus an attenuation of the level of gene knockdown is observed. In support of this hypothesis, cells transfected with pools of functional siRNA — identified by rational design algorithms — consistently generate strong silencing properties, increasing the frequency in which extremely potent silencing (>95%) can be achieved from 48% (with single functional siRNA) to 85% (with pools of siRNA). In addition, these reagents exhibit prolonged silencing effects (e.g., enhanced longevity), suggesting that rational design and pooling can significantly improve the degree of silencing over reagents chosen by random or conventional design parameters.

Delivery

Like all pharmaceutical reagents, delivery of rationally designed siRNA to the proper tissues must overcome a number of hurdles. Issues pertaining to half-life, uptake, knockdown longevity and off-target effects challenge future applications of siRNA as therapeutic agents. These obstacles are poignant especially for nucleic acid pharmaceuticals, where serum stability of unmodified molecules can be measured in tens of seconds.

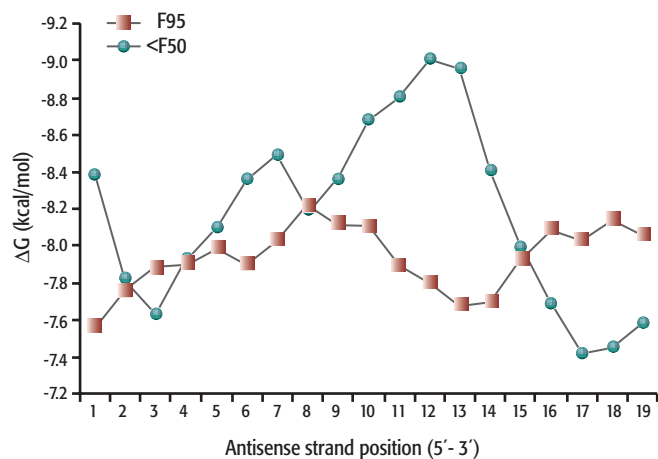


Figure 4. A comparison of the internal stability profile for highly functional (F95) and non-functional (<F50) duplexes. The X-axis indicates the nucleotide position for the sense strand. The Y-axis indicates the relative stability (in kcal/mol) of each position, as determined using nearest-neighbor thermodynamic parameters. The Y-axis represents the nucleotide position along the antisense strand. Boxes indicate values for F95 duplexes. Circles represent the values for non-functional (<F50) duplexes.

Delivery of siRNAs currently is being addressed using a variety of approaches. Plasmid and viral vector delivery systems designed to express siRNA *in situ* — either as duplex molecules or monomeric short hairpin structures (shRNA) — are being tested in a number of laboratories. The challenges of these approaches clearly overlap the trials experienced by those working in the field of gene therapy and include tissue-specific delivery, consistently regulated expression, risks associated with disruption of critical host genes via integration and toxicity. For the expression of duplexes, intracellular transcription of small RNA molecules has been achieved by cloning siRNA-encoding templates into a cassette that contains an RNA polymerase III promoter (typically H1 or U6) upstream of the sequence that codes for the sense or antisense structure, followed by an RNA polymerase termination signal. Subsequent expression of sense and antisense strands from separate vectors (29, 30) involves several steps, most importantly the formation of the functional duplexes by intracellular annealing. From the standpoint of efficiency and therapeutic efficacy, this mode of inducing long-term expression of siRNA is questionable. The presence of a large number of intracellular RNases requires that high concentrations of both strands be expressed in order to generate functional concentrations of the siRNA. Such elevated levels potentially are toxic, possibly inducing cell death through the previously discussed interferon response pathway. Equally important is the fact that generation of intact siRNA from individual sense and antisense strands is unattractive from a therapeutic standpoint. Current FDA requirements necessitate that all entities intended for therapeutic usage must be tested thoroughly. Thus, in the case of a bipartite molecule such as an siRNA, all three structures (the sense strand, the antisense strand and the duplex) would need to be screened for potential side effects, greatly increasing the costs of development. One alternative to this ap-

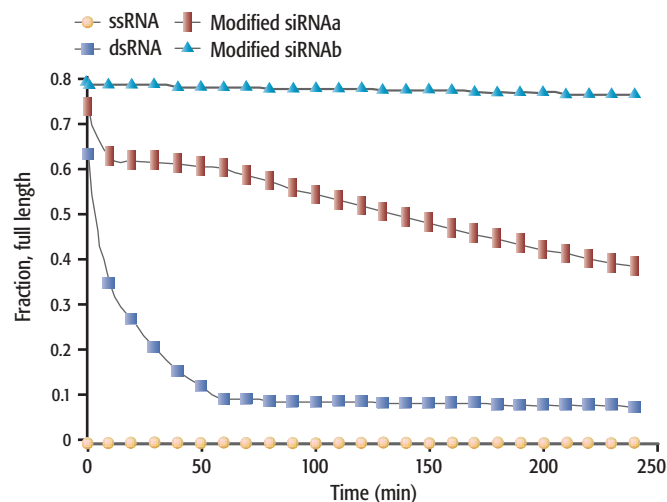


Figure 5. A comparison of the relative stability of a single-stranded (ssRNA, circles), double-stranded (dsRNA, squares) and two chemically modified duplexes (siRNAa, rectangles and siRNAb, triangles) in human serum. Samples of each were incubated in 100% serum at 37 °C, and aliquots were removed at each time point and assessed by polyacrylamide gel electrophoresis (PAGE) analysis.

proach involves expression of siRNAs as single-stranded hairpin structures (31–35). The concept of using shRNAs as a gene silencing tool initially arose from investigations of a distinct class of small RNAs known as microRNAs (miRNAs) or small temporal RNAs (stRNAs) (36, 37) that appear to function in an RNAi-related pathway that regulates gene expression via translational inhibition. miRNA precursors are predicted to form large stem-loop structures that are processed initially by Drosha (an RNase III endonuclease) in the nucleus. The pre-miRNAs then are exported to the cytoplasm, where they are processed by Dicer into small, functional, single-stranded intermediates with inhibitory activity (38, 39). Studies have revealed that shRNAs, modeled on theoretical miRNA structures, are functional and dependent on Dicer activity (39). Thus, shRNAs appear to serve as precursors of functional siRNA silencing.

While shRNAs are attractive, uni-molecular entities, the dependence on Dicer processing adds an additional level of complexity to gene silencing strategies. Several studies designed to analyze the structural requirements of functional shRNA have revealed that the length of the duplex contributes to shRNA functionality (40), as does the presence, size, sequence and position of both stem and loop structures. This is supported by comparisons of chemically synthesized siRNAs and shRNA, where significant variability in silencing efficiencies were observed for structures containing divergent shRNA organizations. siRNAs that routinely induced high levels of gene silencing (90–95% reduction in transcript levels) were diluted significantly (40–80% silencing) when the same sequence was introduced in a hairpin format. Thus the key structural and thermodynamic properties necessary for efficient conversion of shRNA into functional siRNA still are uncertain.

The approach taken to optimize delivery of chemically synthesized siRNAs has focused on incorporating one or more stabilizing groups to functional siRNA. While 2'-O-ACE chemistry is uniquely qualified to tackle this challenge, the addition of chemical modifications introduces additional hurdles in that the newly modified molecules must retain the potency of their unmodified counterparts, exhibit minimal toxicity and be amenable to high-throughput procedures. Progress in this approach already has been made on several fronts. Recently, Dharmacon (Lafayette, Colorado, USA) identified two novel modification strategies (applicable to any sequence) that enhance the serum stability of siRNAs without altering functionality (Figure 5a, 5b). The first of these modifications increases the serum stability of the siRNA, from seconds to more than five hours, without added toxicity. The second modification extends serum stability to over five days and exhibits only minor toxic effects. Again, both of these molecules retain almost full potency. As additional modifications have been developed that eliminate sense strand off-target effects, the reality of developing modified synthetic siRNA into therapeutic reagents is not unreasonable.

In conclusion, the identification of the RNAi pathway will make significant contributions to basic research and medical fields. Broad and rapid application of this technology will demand identification of siRNA design and delivery standards that will ensure duplex functionality in cell culture and whole animal studies. Rational design algorithms, such as those developed by Dharmacon and others, will significantly improve duplex si-

lencing efficiency. Similarly, future development of siRNA chemical modifications and si/shRNA expression systems should extend this technology into new therapeutic arenas.

References

1. L.A. Herzberg et al., *Clin. Chem.* **48(10)**, 1819–1827 (2002).
2. K. Carr, *Nature* **365(6448)**, 685 (1993).
3. M. Minsky, *Scanning* **10(4)**, 128–138 (1998).
4. S. Fields and O. Song, *Nature* **340(6230)**, 245–246 (1989).
5. C. Napoli, C. Lemieux and R. Jorgensen, *Plant Cell* **2(4)**, 279–289 (1990).
6. V. Chandler, W. Eggleston and J. Dorweiler, *Plant Mol. Biol.* **43(2–3)**, 121–145 (2000).
7. N. Romano and G. Macino, *Mol. Microbiol.* **6(22)**, 3343–3353 (1992).
8. A. Fire et al., *Nature* **391(6669)**, 806–811 (1998).
9. R.F. Ketting et al., *Genes & Dev.* **15(20)**, 2654–2659 (2001).
10. T. Tuschl et al., *Genes & Development* **13(24)**, 3191–3197 (1999).
11. P.D. Zamore et al., *Cell* **101(1)**, 25–33 (2000).
12. R.H. Plasterk, *Science* **296(5571)**, 1263–1265 (2002).
13. S.M. Elbashir et al., *Nature* **411(6836)**, 494–498 (2001).
14. M.H. Caruthers, *Science* **230(4723)**, 281–285 (1985).
15. N. Usman et al., *Journal of the American Chemical Society* **109, 7845** (1987).
16. X. Wu and S. Pitsch, *Nucleic Acids Res.* **26(19)**, 4315–4323 (1998).
17. S.A. Scaringe, *Methods in Enzymology* **317**, 3–18 (2000).
18. A.L. Jackson et al., *Nature Biotechnology* **21(6)**, 635–637 (2003).
19. D. Semizarov et al., *Proc. Natl. Acad. Sci. USA* **100(11)**, 6347–6352 (2003).
20. A. Reynolds et al., *Nature Biotechnology* **22(2)**, 326–330 (2004).
21. A. Khvorova, A. Reynolds and S. Jayasena, *Cell* **115(1)**, 209–216 (2003).
22. D.S. Schwarz et al., *Cell* **115(1)**, 199–208 (2003).
23. C. Llave et al., *Plant Cell* **14(7)**, 1605–1619 (2002).
24. T. Holen et al., *Nucleic Acids Res.* **31(9)**, 2401–2407 (2003).
25. J. Martinez et al., *Cell* **110(5)**, 563–574 (2002).
26. M.T. McManus et al., *J. Immunol.* **169(10)**, 5754–5760 (2002).
27. P. Gonczy et al., *Nature* **408(6810)**, 331–336 (2000).
28. C. Llave et al., *Science* **297(5589)**, 2053–2056 (2002).
29. J-Y. Yu, S.L. DeRuiter and D.L. Turner, *PNAS* **092143499** (2002).
30. M. Miyagishi and K. Taira, *Nature Biotechnol.* **20(5)**, 497–500 (2002).
31. T.R. Brummelkamp, R. Bernards and R. Agami, *Science* **296(5567)**, 550–553 (2002).
32. C.P. Paul et al., *Nature Biotechnol.* **20(5)**, 505–508 (2002).
33. P.J. Paddison et al., *Genes & Dev.* **16(8)**, 948–941 (2002).
34. G. Sui et al., *Proc. Natl. Acad. Sci. USA* **99(8)**, 5515–5520 (2002).
35. J-Y. Yu, *Molecular Therapy* **7(2)**, 228 (2003).
36. Y. Zeng, R. Yi and B.R. Cullen, *PNAS* **100(17)**, 9779–9784 (2003).
37. A. Grishok et al., *Cell* **106(1)**, 23–34 (2001).
38. Y. Lee et al., *Nature* **425**, 415–419 (2003).
39. M.T. McManus et al., *RNA* **8(6)**, 842–850 (2002).
40. Y. Zeng, *RNA* **9**, 112–123 (2003). **PG**

Jon Karpilow and **Devin Leake** are staff scientists and **Bill Marshall*** is executive vice president of research and production at Dharmacon Inc. Bill Marshall can be reached at Dharmacon Inc., 2650 Crescent Drive, #100, Lafayette, Colorado 80026 USA. Tel. 303-604-9499, e-mail Marshall.B@Dharmacon.com.

*To whom all correspondence should be addressed.