

# Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity

AIMEE L. JACKSON, JULJA BURCHARD, JANELL SCHELTER, B. NELSON CHAU, MICHELE CLEARY, LEE LIM, and PETER S. LINSLEY

Rosetta Inpharmatics, LLC, a wholly owned subsidiary of Merck & Co., Inc., Seattle, Washington 98109, USA

## ABSTRACT

Transfected siRNAs and miRNAs regulate numerous transcripts that have only limited complementarity to the active strand of the RNA duplex. This process reflects natural target regulation by miRNAs, but is an unintended (“off-target”) consequence of siRNA-mediated silencing. Here we demonstrate that this unintended off-target silencing is widespread, and occurs in a manner reminiscent of target silencing by miRNAs. A high proportion of unintended transcripts silenced by siRNAs showed 3′ UTR sequence complementarity to the seed region of the siRNA. Base mismatches within the siRNA seed region reduced the set of original off-target transcripts but generated new sets of silenced transcripts with sequence complementarity to the mismatched seed sequence. An inducible shRNA silenced a subset of transcripts that were silenced by an siRNA of the same sequence, demonstrating that unintended silencing is sequence mediated and is independent of delivery method. In all cases, off-target transcript silencing was accompanied by loss of the corresponding protein and occurred with dependence on siRNA concentration similar to that of silencing of the target transcript. Thus, short stretches of sequence complementarity to the siRNA or shRNA seed region are key to the silencing of unintended transcripts.

**Keywords:** siRNA; RNAi; seed region; off-target; miRNA; shRNA

## INTRODUCTION

RNA interference (RNAi) is widely used for functional genetic studies and genetic screens in a number of organisms. However, recent investigations have shown that siRNA-mediated silencing may be less specific than was originally believed. Microarray profiling experiments have demonstrated siRNA-mediated silencing of numerous unintended (off-target) transcripts (Jackson et al. 2003; Semizarov et al. 2003; Scacheri et al. 2004). Sequence analysis of several off-target transcripts revealed only partial complementarity with the transfected siRNAs, notably at the 5′ end of the siRNA guide strand (Jackson et al. 2003). In a subsequent study, several false-positive hits resulting from an siRNA screen showed sequence complementarity with the 5′ end of the siRNA guide strand (Lin et al. 2005). This portion of the guide strand is analogous to the “seed” region at the 5′ end of a microRNA that is implicated in

gene regulation by inhibition of translation and/or mRNA degradation (Lai 2002; Lewis et al. 2003; Stark et al. 2003; Yekta et al. 2004; Bagga et al. 2005). Mechanistic studies demonstrated that perfect complementarity with a specific target at the 5′ end of the siRNA antisense strand was sufficient to direct siRNA-target binding and cleavage (Haley and Zamore 2004).

The ability of siRNAs to regulate large numbers of transcripts is similar to the regulation of large numbers of transcripts by miRNAs (Lewis et al. 2003; Stark et al. 2003; Brennecke et al. 2005; Farh et al. 2005; Grun et al. 2005; Krek et al. 2005; Lim et al. 2005; Xie et al. 2005). Intended and unintended transcript silencing could not be distinguished by reducing the siRNA concentration or time of transfection (Jackson et al. 2003). Likewise, on-target and off-target hits in phenotypic screens could also not be distinguished by decreasing the siRNA concentration (Lin et al. 2005). Taken together, these findings argue that off-target transcript silencing is siRNA sequence dependent and is a fundamental feature of RNAi-mediated silencing.

Previously, we demonstrated sequence-dependent off-target activity for a small number of siRNAs. Here, we

**Reprint requests to:** Aimee L. Jackson, Rosetta Inpharmatics, LLC, 401 Terry Avenue N, Seattle, WA 98109, USA; e-mail: [aimee\\_jackson@merck.com](mailto:aimee_jackson@merck.com); fax: (206) 802-6388.

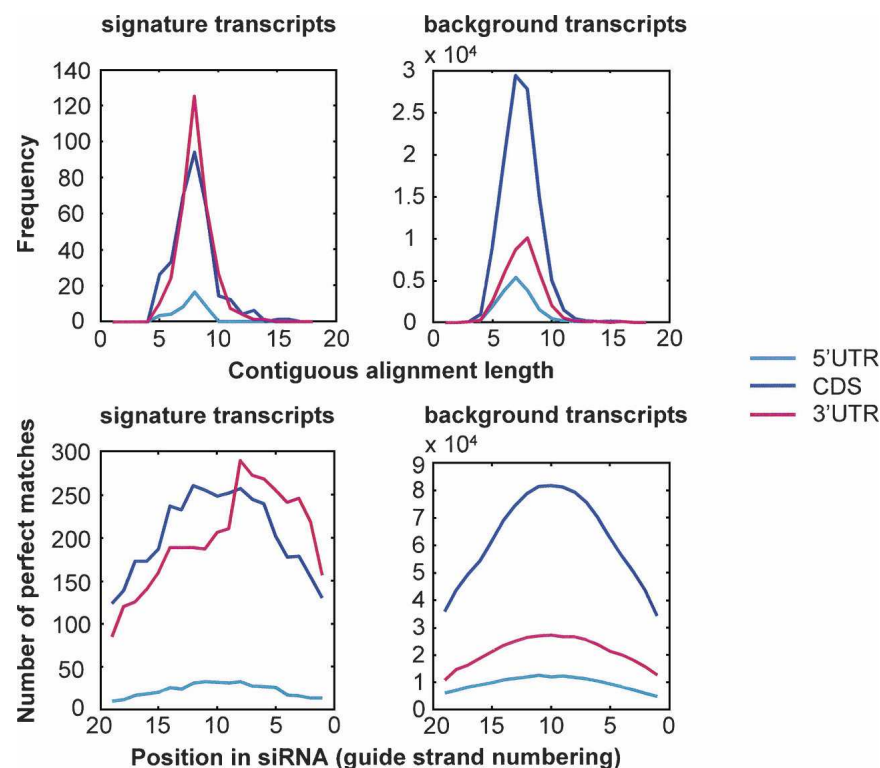
Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.25706>.

extend these findings to a larger number of siRNAs targeting different and unrelated transcripts. We demonstrate that sequence complementarity to the seed region of the siRNA is a primary determinant of regulation of unintended transcripts and their corresponding proteins by all siRNAs examined. We extend this finding to expressed shRNAs, demonstrating that sequence-dependent off-target transcript regulation is independent of delivery method and occurs in multiple cell types.

## RESULTS

Previously, we demonstrated that some off-target transcripts were complementary to a transfected siRNA only in the seed region of the guide strand, whereas others showed longer (11–13 nt) complementarity centrally located in the siRNA sequence (Jackson *et al.* 2003). To broaden our data set, we designed six siRNAs each to six different genes, transfected those duplexes into HeLa cells, and examined transcript regulation by microarray analysis. Signature transcripts (regulated with a  $p$ -value  $\leq 0.01$  at 12 and 24 h post-transfection) for each siRNA were analyzed by FASTA for sequence alignment to the siRNA guide strand. In total, >969 off-target transcripts were analyzed. The longest contiguous regions within each sequence alignment had an average length of  $\sim 8$  nt and were found most frequently in the 3' UTR, CDS, and 5' UTR, respectively (Fig. 1). For comparison, we analyzed siRNA sequence alignment to a background set of transcripts, representing all of the transcripts on the microarray with a defined 3' UTR (10,631 transcripts). The average contiguous alignment for background transcripts was slightly shorter ( $\sim 7$  nt) than for signature transcripts. Although the difference in alignment length between signature and background transcripts is small, it is statistically significant ( $p$ -value  $< 5 \times 10^{-28}$ , Kolmogorov–Smirnov goodness-of-fit hypothesis test) and may contribute to the preferential silencing of these transcripts. In addition, the 3' UTRs of background transcripts showed proportionally fewer guide strand alignments than the signature transcripts (hypergeometric  $p$ -value = 0). Thus, siRNA-regulated transcripts show both quantitative and qualitative differences in how they align with the guide strands.

We next examined which siRNA positions share sequence complementarity with off-target transcripts. We used FASTA alignment to score for the number of perfect sequence matches at each position of the siRNA across the transcripts in the signature set. Signature transcripts with alignments in the 3' UTR were significantly more likely than background transcripts to show bias in alignments toward complementarity with the 5' end of the guide strand (hypergeometric  $p$ -value =  $6 \times 10^{-10}$ ; Fig. 1). In contrast, CDS and 5' UTR alignments of both signature and background transcripts showed a similar preference for the central region of the siRNA, consistent with random matching. These data demonstrate that silencing of unintended transcripts by siRNAs is sequence-based, with a bias for siRNA seed region complementarity to the 3' UTR. This type of target recognition is reminiscent of miRNAs, which regulate hundreds of transcripts through



**FIGURE 1.** siRNA seed regions preferentially align with the 3' UTR regions of unintended transcripts. (*Upper panels*) 36 siRNAs were aligned by FASTA to their down-regulated off-target signature transcripts and to a background set of 10,631 transcript sequences with mapped CDS. The longest contiguous stretch of identity and the transcript region targeted in each guide-strand alignment were determined. Shown on the x-axis is the length of contiguous sequence alignment between the siRNAs and the transcripts they silence. Shown on the y-axis is the frequency within the signature set (*left panel*) or background set (*right panel*) of alignments of the length indicated on the x-axis. (*Bottom panels*) The positions of sequence complementarity between the siRNA guide strand and the target transcript were determined by FASTA alignment. Shown on the x-axis is the position in the siRNA guide strand, with 1 indicating the first position at the 5' end of the guide strand. Shown on the y-axis is the number of perfect matches at each position of the siRNA to transcripts in the signature set (*left panel*) or background set (*right panel*).

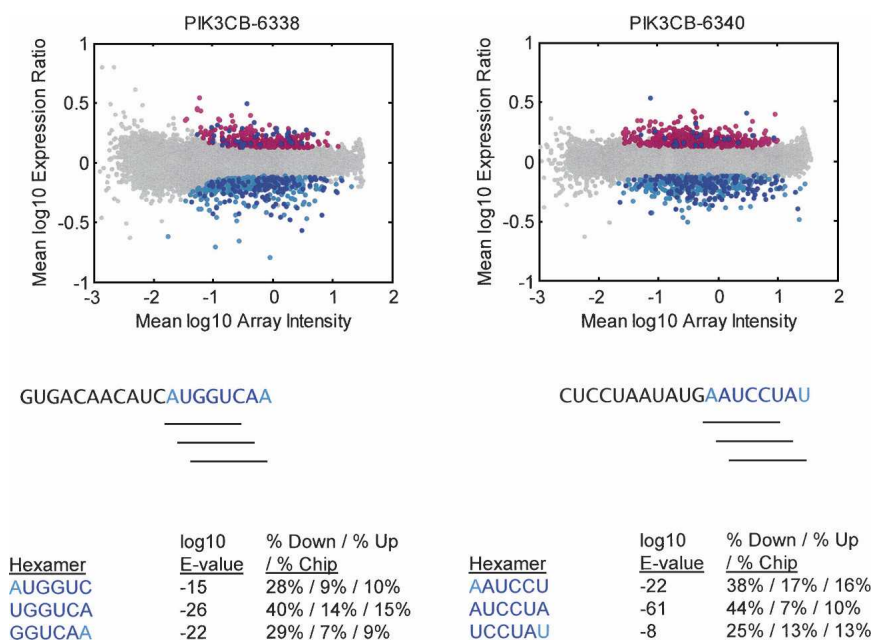
binding of the 7-nt seed of the guide strand to 3' UTRs of target transcripts.

Hexamer motifs corresponding to the seed region of miRNAs are enriched within the 3' UTRs of mRNAs down-regulated by miRNAs (Lim et al. 2005). We hypothesized that sequences complementary to the seed region are also enriched in siRNA off-target transcripts. To test this possibility, we identified the most enriched hexamer sequences in the 3' UTRs of signature transcripts for each siRNA (down-regulated with  $p$ -value  $\leq 0.01$  at 24 h post-transfection). In all cases, the most-enriched hexamer was exactly complementary to positions 2–7 of the guide strand of the siRNA (Fig. 2; data not shown). This finding is in perfect agreement with the findings for hexamer motifs in transcripts regulated by miRNAs (Lim et al. 2005). Hexamers complementary to positions 3–8 or 1–6 of the siRNA guide strand were also enriched in down-regulated transcripts, extending the complementarity to the entire seed region. Two siRNAs of different sequences targeting the same transcript regulated different sets of transcripts, each enriched for those with 3' UTRs complementary to the siRNA

seed region. Thus, siRNA transcript silencing is sequence specific rather than target specific. Examination of the signature transcripts down-regulated with the same kinetics as the intended target in our previous study (MAPK14–193; Jackson et al. 2003), showed that most of these transcripts also had 3' UTRs that contained the MAPK14–193 position 2–7 seed hexamer. The identification of similar transcripts by both FASTA alignment and hexamer motif searching argues that these transcripts are direct siRNA targets.

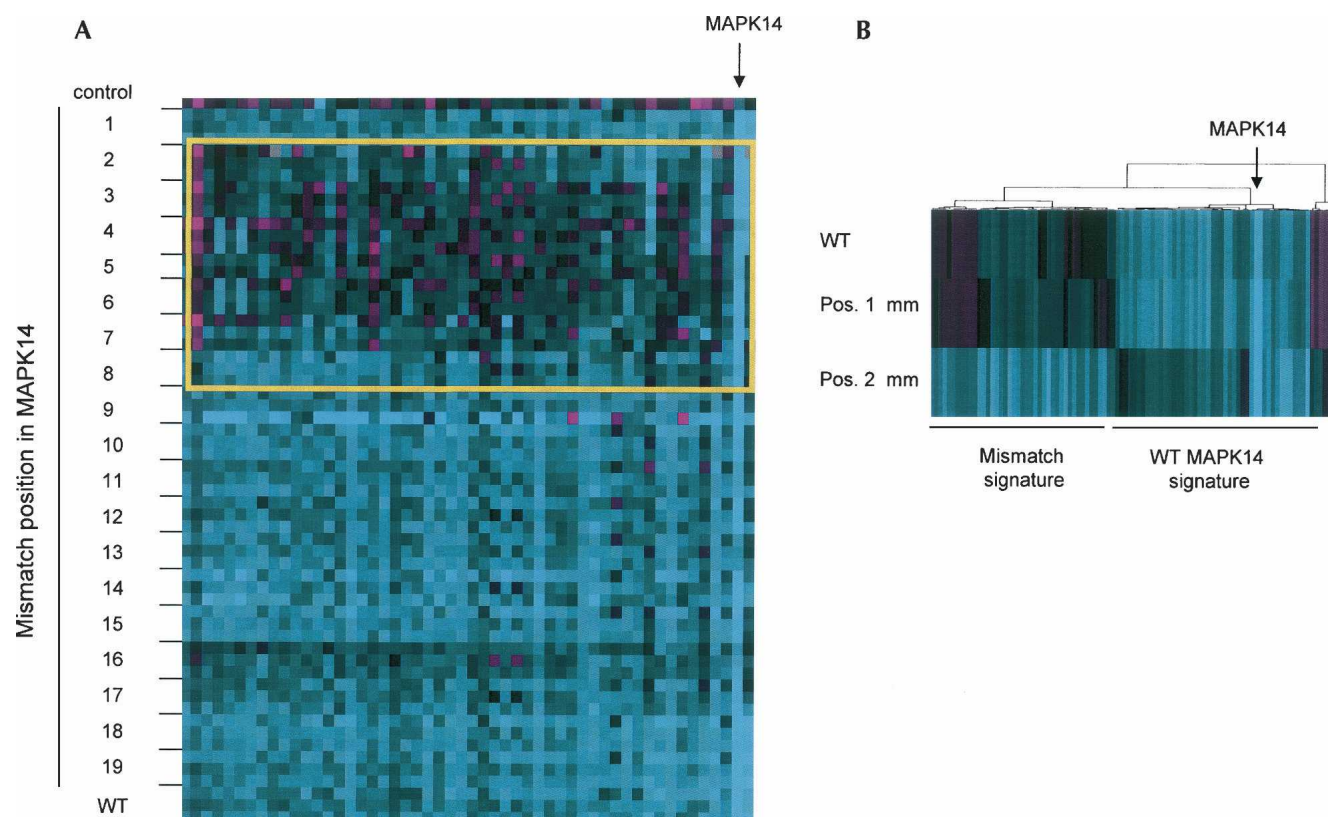
If seed region sequence complementarity is a primary determinant of off-target transcript regulation, then base mismatches within the seed region that disrupt interaction between siRNA and transcript should alter off-target silencing. Some mismatches in the seed region of miRNAs reduced transcript silencing (Doench and Sharp 2004; Lim et al. 2005). To test the effect of mismatches on off-target silencing, we performed a mismatch walk on a siRNA targeting MAPK14. We designed a series of siRNAs that contained mismatches at each individual position of the original sequence. To eliminate effects of bulges on duplex unwinding, residues on both strands of the duplex were

changed to preserve base pairing. We transfected these duplexes into HeLa cells and measured changes in transcript regulation by microarray analysis. The off-target transcript signature of the wild-type MAPK14 siRNA was generally reduced by mismatches introduced into positions 2–7 of the siRNA seed region (Fig. 3A), confirming the importance of this region for off-target silencing. In contrast, silencing of the intended target MAPK14 was not affected by any single mismatch in the seed region. Single base mismatches at positions distal to the seed region had minimal effect on the off-target signature or on the silencing of MAPK14. siRNAs containing seed region mismatches generated a new off-target signature (Fig. 3B). This new off-target signature was complementary to the altered siRNA sequence, as determined by FASTA alignment (data not shown). Thus, the use of mismatched siRNAs may reduce off-target transcript regulations, but this benefit is offset by the concomitant appearance of new off-target transcripts with complementarity to the mismatched seed region. Therefore, seed region sequence complementarity may not be essential for otherwise perfectly matched targets, but is a primary determinant of off-target transcript silencing.



**FIGURE 2.** Signature transcript 3' UTRs contain sequence complementarity to siRNA seed regions. The indicated siRNAs were transfected into HeLa cells. Gene expression signatures ( $p$ -value  $\leq 0.01$ ) for each siRNA were determined 24 h post-transfection. Down-regulated transcripts are indicated in light blue, up-regulated transcripts are indicated in magenta. Gray indicates no significant change in transcript level in response to the siRNA. The sense sequence for each siRNA is indicated, with nucleotides complementary to positions 2–7 of the seed region of the guide strand indicated in dark blue. Positions 1 and 8 of the guide strand are indicated in light blue to highlight the entire seed region of the siRNA. The lines below the sequences indicate the positions of the three seed region hexamers. The 3' UTRs of signature transcripts were searched for complementarity to siRNA seed region hexamers. Shown are the  $e$ -values ( $p$ -value with Bonferroni correction) for down-regulated transcripts with 3' UTR complementarity to each seed region hexamer. Also shown is the percentage of down-regulated, up-regulated, or background transcripts containing 3' UTR complementarity to the hexamer.



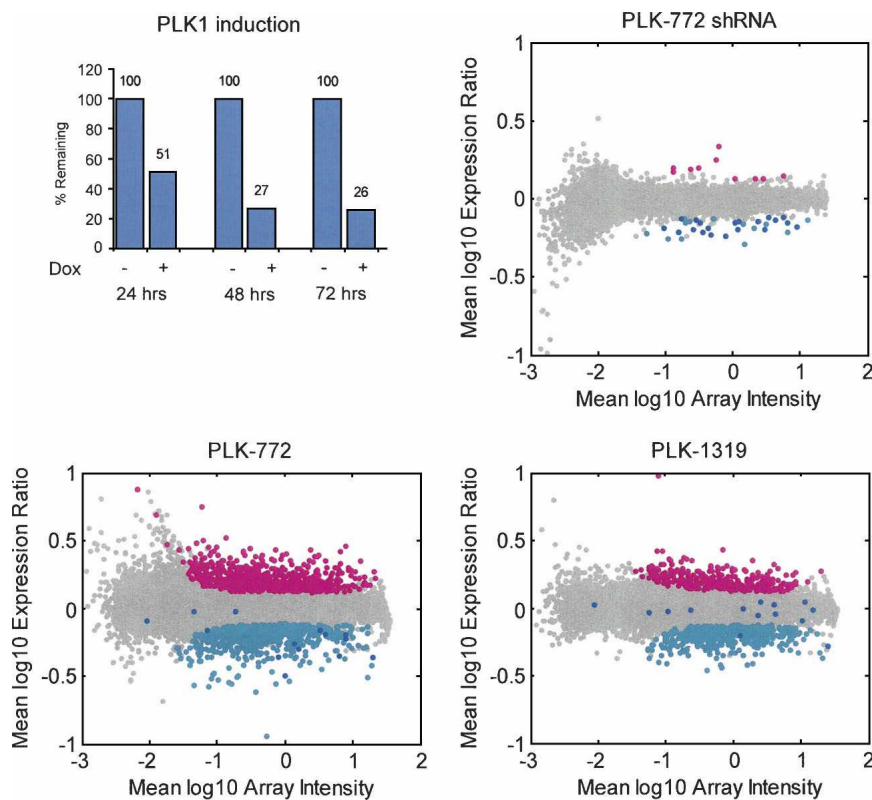


**FIGURE 3.** Position-specific effect of base mismatches on silencing of off-target transcripts. (A) siRNAs to MAPK14–193 were synthesized to contain a single base substitution at each position. siRNAs were then transfected into HeLa cells, and changes in transcript regulation were analyzed by microarray profiling (Jackson et al. 2003). Shown is a heatmap representing 52 genes regulated (x-axis) in 61 experiments (y-axis). These 52 transcripts represent the consensus signature of down-regulated transcripts for the MAPK14 siRNA. Transcripts shown were regulated with  $p \leq 0.01$ . No cuts were placed on fold regulation. Transcripts down-regulated in siRNA-transfected cells are shown in light blue, transcripts up-regulated in siRNA-transfected cells are shown in magenta. Black indicates no change in transcript level. Three independent experiments are shown for each base mismatch. The gold box indicates the location of the siRNA seed region (positions 2–8 of the guide strand). Transcripts are ordered by percent change in down-regulation (normalized mean log ratio change) across the wild-type signature. The arrow indicates the location of the target transcript MAPK14. (B) siRNAs to MAPK14–193 containing mismatches at position 1 or position 2 of the seed region were transfected into HeLa cells, and changes in transcript regulation were analyzed by microarray as described in A.

To determine whether seed-region-dependent off-target activity extends to shRNAs, we used an inducible promoter to express an shRNA targeting the essential gene *PLK1*. The shRNA construct was delivered to HT29 cells by use of a lentiviral vector. Expression of the shRNA was induced by addition of doxycycline, and RNA was isolated 24 h post-induction. At this time post-induction, the shRNA produced 51% silencing of *PLK1*. Although *PLK1* silencing was maximal at later times post-induction (Fig. 4), cells were beginning to undergo apoptosis at these times, complicating interpretation of the expression signatures. Transcripts regulated by the shRNA at 24 h were identified by expression profiling of RNA from doxycycline-induced cells compared with untreated cells. Of 33 transcripts down-regulated by the shRNA ( $p$ -value  $\leq 0.01$ ; Fig. 4), 20 (61%) displayed 3' UTR sequence complementarity to positions 1–6, 2–7, and 3–8 hexamers of the shRNA guide strand ( $e$ -values  $7.9 \times 10^{-5}$ ,  $4.4 \times$

$10^{-9}$ , and  $1.4 \times 10^{-6}$ , respectively). This enrichment demonstrates that shRNAs, like siRNAs, regulate transcripts complementary to the seed region of the guide strand.

We examined the regulation of these shRNA off-target transcripts in expression signatures resulting from transfection of HeLa cells with an siRNA of the same sequence as the *PLK1* shRNA (*PLK1*–772). Of the 20 seed region complementary transcripts regulated by the shRNA in HT29 cells, 14 were also present on the microarrays used for HeLa profiling, and 10 of these were expressed at significant levels ( $\log_{10}$  intensity  $\geq -1$ ,  $\sim 1$  copy/cell) in HeLa cells. These 10 shRNA off-target transcripts were also significantly down-regulated ( $p$ -value of regulation  $\leq 0.01$ ) by a *PLK1*–772 siRNA (Fig. 4). Thus, although two different cell lines were used, an induced shRNA and a siRNA of the same sequence share sets of off-target transcripts with seed region complementarity.



**FIGURE 4.** shRNA and siRNA of the same sequence regulate the same subset of seed-region-complementary transcripts. HT29 cells were transduced with lentivirus expressing PLK1 shRNA under control of a doxycycline-inducible promoter. PLK1 transcript silencing at times following induction of the shRNA was measured by TaqMan qPCR (*upper left* panel). The gene expression signature resulting from 24-h induction of PLK1 shRNA is shown (*upper right* panel). Transcripts down-regulated by the shRNA are indicated in light blue; transcripts up-regulated by the shRNA are indicated in magenta. Down-regulated transcripts with 3' UTRs that contain sequence complementarity to the shRNA seed region are indicated as dark blue dots. A PLK1 siRNA of the same sequence as the shRNA was transfected into HeLa cells, and gene expression signatures were generated at 24 h post-transfection (*lower left* panel). The transcripts down-regulated by the PLK1 shRNA were highlighted in the siRNA signature and are indicated as dark blue dots. A PLK1 siRNA of a different sequence was transfected into HeLa cells (*lower right* panel). The transcripts regulated by the PLK1 shRNA are indicated as dark blue dots.

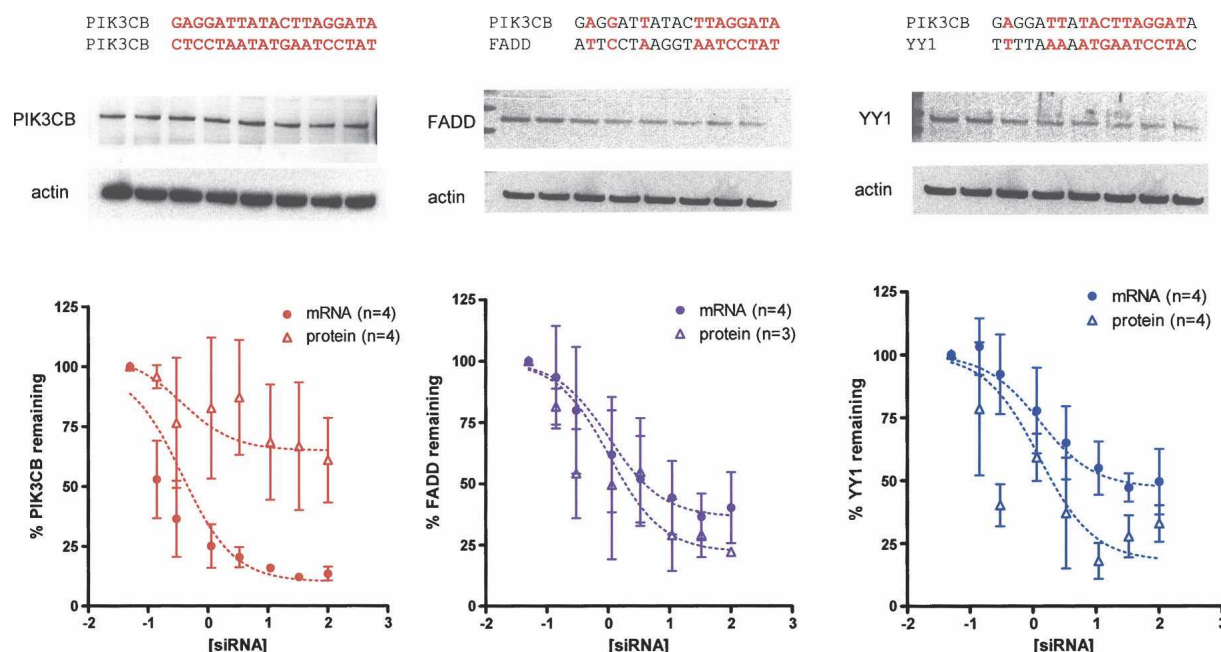
The PLK1-772 shRNA-regulated transcripts were not well regulated by siRNAs targeting different sequences in PLK1. For one of these different siRNAs (PLK1319), only 1 of the 14 shRNA-regulated transcripts expressed on the microarray was significantly down-regulated (Fig. 4). In eight independent experiments with PLK1-772 siRNA, an average of 9 of the 10 expressed shRNA off-target transcripts were also regulated by the siRNA. In contrast, an average of only 1 of the 10 shRNA transcripts was regulated in eight experiments using PLK1 siRNAs of different sequences. The number of shRNA off-target transcripts regulated by siRNAs of the same and different sequences was significantly different ( $p$ -value =  $1.2 \times 10^{-7}$ ,  $t$ -test). The target, PLK1, was silenced in all cases. Thus, this transcript regulation is sequence specific, not target specific, and is independent of the manner in which the regulatory RNA is delivered.

miRNAs regulate their targets primarily at the protein level. If siRNA off-target silencing is related to miRNA activity, then off-target silencing at the protein level will be equally or more pronounced than at the transcript level. To test this possibility, we examined regulation of the proteins corresponding to transcripts with only partial sequence complementarity to the siRNA. From the set of 36 siRNAs described above, we selected siRNAs for which on- and off-target protein products are recognized by commercially available antibodies. siRNA titration analyses rather than single point analyses were used to detect silencing trends and minimize effects of experimental noise. Transcript levels were measured by quantitative RT-PCR analysis 24 h post-transfection, and protein levels were measured by Western analysis 48 h post-transfection. These time points were chosen on the basis of control experiments indicating that RNA silencing is generally maximal  $\sim 24$  h post-transfection and that protein silencing varies depending on the target, but is generally maximal  $\sim 48$ – $72$  h post-transfection.

For the intended target transcripts of the siRNAs, transcript silencing was highly potent ( $IC_{50} < \sim 1$  nM, Fig. 5; data not shown). Silencing of the corresponding protein generally followed a similar dose response as transcript silencing (PIK3CB, Fig. 5; data not shown.) Although the dose response curve for the PIK3CB protein appears shifted to the right, there was no statistically significant difference in  $IC_{50}$

between the protein and RNA silencing. Maximal silencing for the PIK3CB protein is only  $\sim 70\%$  of control at this time point (48 h), possibly reflecting a protein half-life longer than the 48 h assayed. In other experiments, the PIK3CB protein was silenced to 20% of control by 72 h post-transfection (data not shown).

We also examined two off-target transcripts with seed region sequence complementarity to the PIK3CB siRNA. Maximal silencing for these transcripts was only  $\sim 50\%$  of control, less effective than that for the target transcript PIK3CB. Protein silencing showed a dose response that was similar to transcript silencing for both FADD and YY1 (Fig. 5), verifying that the silencing of these off-target transcripts results in changes at the protein level. Notably, FADD and YY1 share only limited complementarity to the PIK3CB siRNA guide strand, predominantly in positions 2–8 of the



**FIGURE 5.** Off-target protein regulation by siRNAs. Decreasing concentrations of an siRNA targeting PIK3CB were transfected into HeLa cells. Silencing of the intended target (PIK3CB) or two off-target transcripts previously identified by microarray profiling (FADD, YY1) was measured by TaqMan analysis at 24 h post-transfection. In parallel samples, protein silencing was measured by Western blotting at 48 h post-transfection. (Top panel) Sequence identity of the siRNA with PIK3CB (on-target), FADD and YY1 (off-target) transcripts. Shown are the sense strand sequence for the siRNA and the transcript sequence for the silenced transcripts. Red font denotes positions of identity. (Middle panel) Western blotting data from a representative experiment. (Bottom panel) Quantitative analysis of mRNA silencing (solid lines) and protein silencing (dashed lines) versus siRNA dose. mRNA levels were normalized to HGUS levels and protein levels were normalized to actin levels. Data are expressed as a percentage of levels in mock-transfected cells. The scatter graphs show means  $\pm$ SD of 3–4 independent experiments.

seed region (Fig. 5). Combining six different comparisons (Fig. 5; data not shown), there was no significant difference in the potency (IC<sub>50</sub>) of on-target or off-target silencing, either at the transcript or protein level ( $p = 0.08$ ,  $t$ -test). Further, there was no significant difference in potency of off-target protein and mRNA silencing ( $p = 0.14$ ,  $t$ -test).

We examined the regulation of signature transcripts and their corresponding protein levels by miRNA duplexes transfected into HeLa cells (data not shown). For all eight miRNAs analyzed, transcript and protein silencing occurred with similar miRNA dose response, analogous to what we found with siRNAs. Taken together, these findings demonstrate that widespread transcript silencing by siRNAs or miRNAs is not simply an artifact of microarray analysis, but predicts measurable changes in protein levels. Conversely, we can predict off-target proteins regulated by siRNAs and protein targets regulated by miRNAs based on transcript regulation and seed region sequence alignment. Furthermore, partial sequence complementarity with the siRNA seed region is important for silencing of unintended transcripts and their corresponding proteins.

## DISCUSSION

When siRNA off-target silencing was first reported (Jackson et al. 2003), it was unclear whether this off-target

activity would impact phenotypic analysis using siRNAs. More recently, however, Lin et al. (2005) reported that all of the top hits from one siRNA screen resulted from off-target silencing. Notably, all of these off-target transcripts shared partial sequence complementarity to the siRNA, consistent with our original results (Jackson et al. 2003). We now demonstrate that the silencing of unintended transcripts by siRNAs is reminiscent of transcript regulation by miRNAs. The seed region of miRNAs plays a key role in target recognition (Wightman et al. 1993; Lai 2002) and is the most highly conserved region of metazoan miRNAs (Lewis et al. 2003; Lim et al. 2003).

Our data suggest that this fundamental role of the seed region in target recognition also extends to siRNAs. Transcripts regulated by siRNAs are enriched for sequences having complementarity to the siRNA seed region. Altering the seed region sequence produces a new off-target signature of transcripts that share sequence complementarity with the new seed sequence. shRNAs also regulate unintended transcripts with seed region complementarity. siRNAs and shRNAs of the same sequence regulate overlapping sets of seed region complementary transcripts, even across cell lines. The regulation of these unintended sequences is independent of the transfection/delivery method. In all cases examined, silencing of the unintended transcripts was accompanied by silencing of the corresponding protein. Thus, these



unintended transcript regulations are a fundamental property of RNAi-mediated silencing.

miRNAs may regulate their targets by translational repression. Since off-target regulation by siRNAs is reminiscent of target regulation by miRNAs, we might expect off-target siRNA regulation to be stronger at the protein level. However, siRNA and miRNA dose–response curves were similar for transcripts and their corresponding proteins. This suggests that neither siRNA off-target nor miRNA target regulations occur preferentially at the protein level. However, this may reflect the fact that we tested a biased set of genes, selected because silencing of their transcripts could be measured by microarray. There are many more potential targets with equivalent seed region sequence complementarity than we are able to detect by microarray. It remains to be determined whether some of these show preferential regulation at the protein level.

Not all transcripts with ~7 nt of sequence complementarity are silenced by siRNAs and shRNAs. One explanation for this finding is that regulated transcripts have certain characteristics that make them preferential targets for the siRNAs. This is similar to the situation with targets of mammalian miRNAs (Farh et al. 2005; Stark et al. 2005). miRNAs regulate hundreds of transcripts through binding of the ~7 nt seed region to target 3' UTRs, but not all transcripts with 7 nt of sequence complementarity are miRNA targets. Functional miRNA targets contain additional specificity determinants. Transcripts are under evolutionary selection for regulation by miRNAs (targets) or to actively avoid regulation by miRNAs (anti-targets). Genes involved in basic, essential cellular processes avoid target sites for all miRNAs (Stark et al. 2005). miRNA targets are well expressed and contain longer 3' UTR regions, while anti-targets have shorter 3' UTRs, which limits miRNA target site density. miRNA anti-targets may also be expressed temporally or spatially to avoid coexpression of the transcript and miRNA.

Transcripts regulated by exogenous siRNAs have not had the benefit of coevolution, yet seem to contain several of the characteristics of miRNA targets (A.L. Jackson, J. Burchard, and P.S. Linsley, unpubl.). Like miRNA targets, we find that siRNA off-target transcripts tend to be well expressed relative to a background set. siRNA off-target transcripts also tend to contain longer 3' UTRs than background transcripts. Importantly, the siRNA off-target transcripts in our studies are enriched for miRNA target sites relative to a background set (J. Burchard and P.S. Linsley, unpubl.). This suggests a tendency for siRNAs to regulate bona fide miRNA targets.

Understanding the characteristics that determine which transcripts will be functional targets for siRNAs may help in the selection of siRNAs predicted to have fewer negative or phenotypic consequences. While improved siRNA design algorithms enhance selection of siRNAs with increased efficacy and potency, there is thus far no indication that

these algorithms yield siRNAs with improved specificity. As no algorithm can eliminate significant numbers of 7–8-base matches of siRNAs to the transcriptome, we predict that perfect specificity will be difficult to achieve. New informatics approaches to address this issue directly may potentially yield more specific siRNAs, but it is unlikely that all unintended activity will be reduced. The regulation of the imperfectly matched targets is a fundamental aspect of siRNA and shRNA biology. Since each siRNA will have a different seed region sequence, and therefore a different complement of unintended transcript regulations, using multiple siRNA duplexes of differing sequence to silence a target of interest remains the surest way to reduce false-positive phenotypes in phenotypic analyses.

## MATERIALS AND METHODS

### Design and synthesis of siRNA duplexes

All siRNAs were designed by Rosetta using a proprietary algorithm and were synthesized by Dharmacon using a modified version of ACE chemistry (Scaringe 2001). See Table 1 for siRNA sequences.

### Microarray analysis

HeLa cells were transfected in 6-well plates using Oligofectamine (Invitrogen) and the indicated doses of siRNA duplex. Where not specified, the concentration of siRNA was 100 nM. RNA was isolated 24 h following transfection. RNA from siRNA-transfected cells was hybridized against RNA from mock-transfected cells (treated with transfection reagent in the absence of RNA duplex). Total RNA was purified by a Qiagen RNeasy kit and processed as described previously (Hughes et al. 2001) for hybridization to microarrays containing oligonucleotides corresponding to ~21,000 human genes. Ratio hybridizations were performed with fluorescent label reversal to eliminate dye bias. Error models have been described previously (Hughes et al. 2001). Data shown are signature genes that display a difference in expression level ( $p < 0.01$ ) relative to mock-transfected cells. No cuts were placed on fold change in expression. Data were analyzed using Rosetta Resolver or MatLab software.

### Sequence alignment using FASTA

Off-target transcripts ( $p < 0.01$ ) with mapped CDS were selected as having sufficiently complete sequence available for alignment. All transcripts with mapped CDS on the microarray used for expression profiling were selected as the background set.

Each strand of the siRNA or miRNA duplex was aligned with each off-target transcript using FASTA 3.4 (Pearson and Lipman 1988) with parameters:

ktup: 2

display up to 50,000 alignments and alignment scores

gap opening penalty: –10

gap extension penalty: –10

match reward/mismatch penalty: +5/–5'

FASTA selects one alignment per siRNA strand/transcript pair.

**TABLE 1.** siRNA sequences used in this study

siRNA	Sense sequence
MAPK14-193	CCUACAGAGAACUGCGGUU
MPHOSPH1-202	GACAUGCGAAUGACACUAG
MPHOSPH1-2692	AUGAAGGAGAGUGAUCACC
PIK3CB-6338	GUGACAACAUCAUGGUCAA
PIK3CB-6340	CUCCUAAUAUGAAUCCUUAU
SOS1-1582	AUUGACCACCAGGUUUCUG
MAPK14-pos1 mismatch	CCUACAGAGAACUGCGGUA
MAPK14-pos2 mismatch	CCUACAGAGAACUGCGGAU
MAPK14-pos1 and 2 mismatch	CCUACAGAGAACUGCGGAA
MAPK14-pos3 mismatch	CCUACAGAGAACUGCGAUU
MAPK14-pos4 mismatch	CCUACAGAGAACUGCAGUU
MAPK14-pos5 mismatch	CCUACAGAGAACUGAGGUU
MAPK14-pos6 mismatch	CCUACAGAGAACUACGGUU
MAPK14-pos7 mismatch	CCUACAGAGAACGGCGGUU
MAPK14-pos8 mismatch	CCUACAGAGAAUUGCGGUU
MAPK14-pos9 mismatch	CCUACAGAGAGCUGCGGUU
MAPK14-pos10 mismatch	CCUACAGAGGACUGCGGUU
MAPK14-pos11 mismatch	CCUACAGAAAACUGCGGUU
MAPK14-pos12 mismatch	CCUACAGGGAACUGCGGUU
MAPK14-pos13 mismatch	CCUACAAAGAACUGCGGUU
MAPK14-pos14 mismatch	CCUACGGAGAACUGCGGUU
MAPK14-pos15 mismatch	CCUAAAGAGAACUGCGGUU
MAPK14-pos16 mismatch	CCUGCAGAGAACUGCGGUU
MAPK14-pos17 mismatch	CCGACAGAGAACUGCGGUU
MAPK14-pos18 mismatch	CUUACAGAGAACUGCGGUU
MAPK14-pos19 mismatch	ACUACAGAGAACUGCGGUU
PLK1 772	GAGACCTACCTCCGGATCA
PLK1 1319	CACGCCTCATCTCTACAA

Alignments of each transcript to each strand of the siRNA were examined for the longest contiguous stretch of identity, and the difference between these lengths was calculated. The strand with the greater excess length of contiguous identity over background across each siRNA's entire down-regulated signature was considered the preferred strand of the siRNA and used for further analysis. Background was calculated as the average excess weight per gene for either strand over alignments of multiple siRNAs to the background set.

## Hexamer complementarity

The enrichment for transcripts that contain a given hexamer within their 3' UTR was carried out by first tabulating all the human genes with LocusLink IDs that contained the hexamer within their annotated 3' UTR, according to Unigene 161 Hs.seq.uniq. The probability of overlap of the microarray signature with a hexamer set was calculated with the hypergeometric cumulative distribution function using a background set of all the genes on the array that contained LocusLink-annotated 3' UTRs.

## Western blotting

Protein lysates were harvested 48 h following transfection of the siRNAs into HeLa cells. Antibodies were from Upstate Biotechnology (PIK3CB), Santa Cruz (FADD, YY1), or Abcam (Actin). Protein levels were normalized to actin levels at each siRNA concentration.

## shRNA expression

HT29 colorectal carcinoma cells (ATCC) were transduced with lentiviruses expressing tTS (tetracycline-controlled transcriptional repressor; Clontech). Pool-infected cells were selected with G418 and expanded. tTS-expressing HT29 cells were subsequently infected with lentiviruses containing a tetracycline-inducible PLK1 shRNA driven from an H1 promoter. Generation of recombinant lentiviruses and transduction of HT29 cells were performed according to the manufacturer's protocols (Invitrogen). For induction of shRNA expression, cells were incubated in the presence of 50 ng/mL of doxycycline, and RNA samples were harvested at indicated time points. PLK1 message levels were analyzed by Taqman qPCR.

## ACKNOWLEDGMENTS

We thank members of the Rosetta biology team for advice and critical reading of the manuscript, Rosetta Gene Expression Lab for microarray hybridizations.

Received January 18, 2006; accepted March 16, 2006.

## REFERENCES

- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. 2005. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122**: 553–563.
- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. 2005. Principles of microRNA-target recognition. *PLoS Biol.* **3**: e85.
- Doench, J.G. and Sharp, P.A. 2004. Specificity of microRNA target selection in translational repression. *Genes & Dev.* **18**: 504–511.
- Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B., and Bartel, D.P. 2005. The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science* **310**: 1817–1821.
- Grun, D., Wang, Y.L., Langenberger, D., Gunsalus, K.C., and Rajewsky, N. 2005. microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comput. Biol.* **1**: e13.
- Haley, B. and Zamore, P.D. 2004. Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.* **11**: 599–606.
- Hughes, T.R., Mao, M., Jones, A.R., Burchard, J., Marton, M.J., Shannon, K.W., Lefkowitz, S.M., Ziman, M., Schelter, J.M., Meyer, M.R., et al. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19**: 342–347.
- Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21**: 635–637.
- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., et al. 2005. Combinatorial microRNA target predictions. *Nat. Genet.* **37**: 495–500.
- Lai, E.C. 2002. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* **30**: 363–364.
- Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. 2003. Prediction of mammalian microRNA targets. *Cell* **115**: 787–798.
- Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. 2003. The microRNAs of *Caenorhabditis elegans*. *Genes & Dev.* **17**: 991–1008.



- Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**: 769–773.
- Lin, S., Ruan, S., Anderson, M.G., McDowell, J.A., Kroeger, P.E., Fesik, S.W., and Shen, Y. 2005. siRNA-mediated off-target gene silencing triggered by a 7nt complementation. *Nucleic Acids Res.* **33**: 4527–4535.
- Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85**: 2444–2448.
- Scacheri, P.C., Rozenblatt-Rosen, O., Caplen, N.J., Wofsyberg, T.G., Umayam, L., Lee, J.C., Hughes, C.M., Shanmugam, K.S., Bhattacharjee, A., Meyerson, M., et al. 2004. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc. Natl. Acad. Sci.* **101**: 1892–1897.
- Scaringe, S.A. 2001. RNA oligonucleotide synthesis via 5'-silyl-2'-orthoester chemistry. *Methods* **23**: 206–217.
- Semizarov, D., Frost, L., Sarthy, A., Kroeger, P., Halbert, D.N., and Fesik, S.W. 2003. Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl. Acad. Sci.* **100**: 6347–6352.
- Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. 2003. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1**: E60.
- Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. 2005. Animal microRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **123**: 1133–1146.
- Wightman, B., Ha, I., and Ruvkun, G. 1993. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75**: 855–862.
- Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**: 338–345.
- Yekta, S., Shih, I., and Bartel, D.P. 2004. MicroRNA-directed cleavage of *HOXB8* mRNA. *Science* **304**: 594–596.



# RNA

A PUBLICATION OF THE RNA SOCIETY

## Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity

Aimee L. Jackson, Julja Burchard, Janell Schelter, et al.

RNA 2006 12: 1179-1187

---

**References** This article cites 20 articles, 7 of which can be accessed free at:  
<http://rnajournal.cshlp.org/content/12/7/1179.full.html#ref-list-1>

**Open Access** Freely available online through the *RNA* Open Access option.

**License** Freely available online through the open access option.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

Dharmacon<sup>™</sup> Reagents  
Custom synthesis, RNAi, and CRISPR solutions

Infinite  
Reliability

More

horizon  
a PerkinElmer company

---

To subscribe to *RNA* go to:  
<http://rnajournal.cshlp.org/subscriptions>

---