Update TRENDS in Pharmacological Sciences Vol.28 No.4

Research Focus

Off-target effects associated with long dsRNAs in *Drosophila* RNAi screens

Jason Moffat[1,2*], Jan H. Reiling[1,2*] and David M. Sabatini[1,2]

[1] Whitehead Institute, 9 Cambridge Center, Cambridge, MA 02142, USA
[2] Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Evidence of off-target effects (OTEs) associated with small interfering (si)RNAs (19–29 bp) in mammalian cells has existed for several years. Two recent articles demonstrate that short sequences within long double-stranded (ds)RNAs frequently cause undesirable OTEs in cultured *Drosophila* cells. These results reveal the potential for high false-positive rates in RNA interference (RNAi) screens using long dsRNAs and highlight the need for screening with multiple, non-overlapping long dsRNAs or siRNAs. Discovering multiple potent siRNAs with minimal off-target profiles for each target transcript will be invaluable for genome-based studies of gene function and for personalized RNAi therapeutics.

Introduction

RNA interference (RNAi) is a method used to knock-down gene expression by targeting mRNA; within a decade, RNAi has gone from being an interesting biological phenomenon to being a tool that has transformed scientific inquiry and that has the potential to treat disease [1]. RNAi entered the spotlight when it was discovered that the core machinery it requires is conserved across the fungi, plant and animal kingdoms. Because the mechanistic understanding of the RNAi pathway has progressed at such a rapid pace, it is important to adapt and amplify the technology appropriately. A recent wave of perspectives has highlighted some of the important steps that are required to confirm the specificity of RNAi experiments [2,3]. In this article, we highlight some of the main findings related to off-target effects (OTEs; see Glossary) associated with long double-stranded (ds)RNAs in cultured *Drosophila* cells and compare this with what is already known about OTEs in mammalian cells.

The two general approaches for employing RNAi in systematic genome-wide screens are: (i) the delivery of long dsRNAs to cells and animals by transfection and feeding, respectively; and (ii) the use of small interfering (si)RNAs delivered by transfection or viral transduction [4]. Current RNAi protocols for cultured *Drosophila* cells use *in vitro* transcribed dsRNA that is several hundred base pairs in length and can be transfected by lipid-based delivery or serum-starvation-induced entry into cells. Upon entry, long dsRNAs are cleaved into a pool of 21–23-nucleotide (nt) siRNAs by Dicer, a dsRNA-specific RNase III. The siRNAs are then incorporated into the RNA-induced silencing complex (RISC). RISC-induced processing of the loaded siRNA leads to siRNA unwinding, followed by selection of the guide strand and degradation of the passenger strand. The activated RISC complex recognizes the target mRNA by siRNA–mRNA base pairing and, ultimately, cleaves the mRNA through the action of Slicer. Thus, the length and sequence of the siRNA are the keys to obtaining silencing specificity and avoiding OTEs.

RNAi in mammals is most commonly achieved by direct transfection of ~21–29-nt siRNAs or by expression of short hairpin (sh)RNAs that become processed to siRNAs. It is recognized in the mammalian RNAi field that OTEs complicate the analysis and identification of candidate genes that score in a primary screen. Several factors have been proposed to account for OTEs (Figure 1): (i) dsRNA strands >30 nts can induce the interferon (IFN) or Toll-like receptor (TLR) response in mammalian cells [5–9]; (ii) excessive numbers of siRNA can saturate the endogenous RNAi machinery, creating neomorphic phenotypes [10]; and (iii) the guide or passenger strand of the siRNAs can hybridize with unintended nucleic acids. In this final scenario, matches as short as six or seven nts can act like the 'seed' region in a micro (mi)RNA and cause translational repression of several unintended targets [11–13]. The permissiveness of siRNAs in interfering with targets of limited sequence similarity makes predictions about OTEs especially difficult. Curiously, nonspecific RNAi effects have been observed only recently in the fly community and have not been reported as a problem in the *Caenorhabditis elegans* field.

Evidence of OTEs in *Drosophila* cells

To discover new components of the Wingless (Wg) signal transduction pathway, Ma *et al.* performed a high-throughput RNAi screen in cultured *Drosophila*

Glossary

CAN: triplet nucleotide repeat sequences (CAA, CAT, CAC or CAG).
Guide strand: antisense RNA strand of the siRNA duplex that hybridizes to a complementary mRNA sequence and directs the cleavage of that message.
Iterative mapping experiments: the use of successively smaller targeting fragments to define a common region between different dsRNAs that confer a similar phenotype.
Off-target effects (OTEs): effects that arise from the degradation of unintended transcripts, the inadvertent activation of IFN or TLR responses or the perturbation of the endogenous RNAi machinery to prevent the degradation or silencing of natural targets.
Passenger strand: sense RNA strand of siRNA duplex that does not hybridize to target mRNA sequences and is discarded when the active strand is incorporated into the RNAi-induced silencing complex.
TCF-binding sites: nucleotide sequence [(C/T)(C/G)TTTGAT] recognized by TCF–LEF (lymphoid enhancer factor) transcription factor complexes.

Corresponding author: Sabatini, D.M. (sabatini@wi.mit.edu).
* These authors contributed equally.
Available online 9 March 2007.

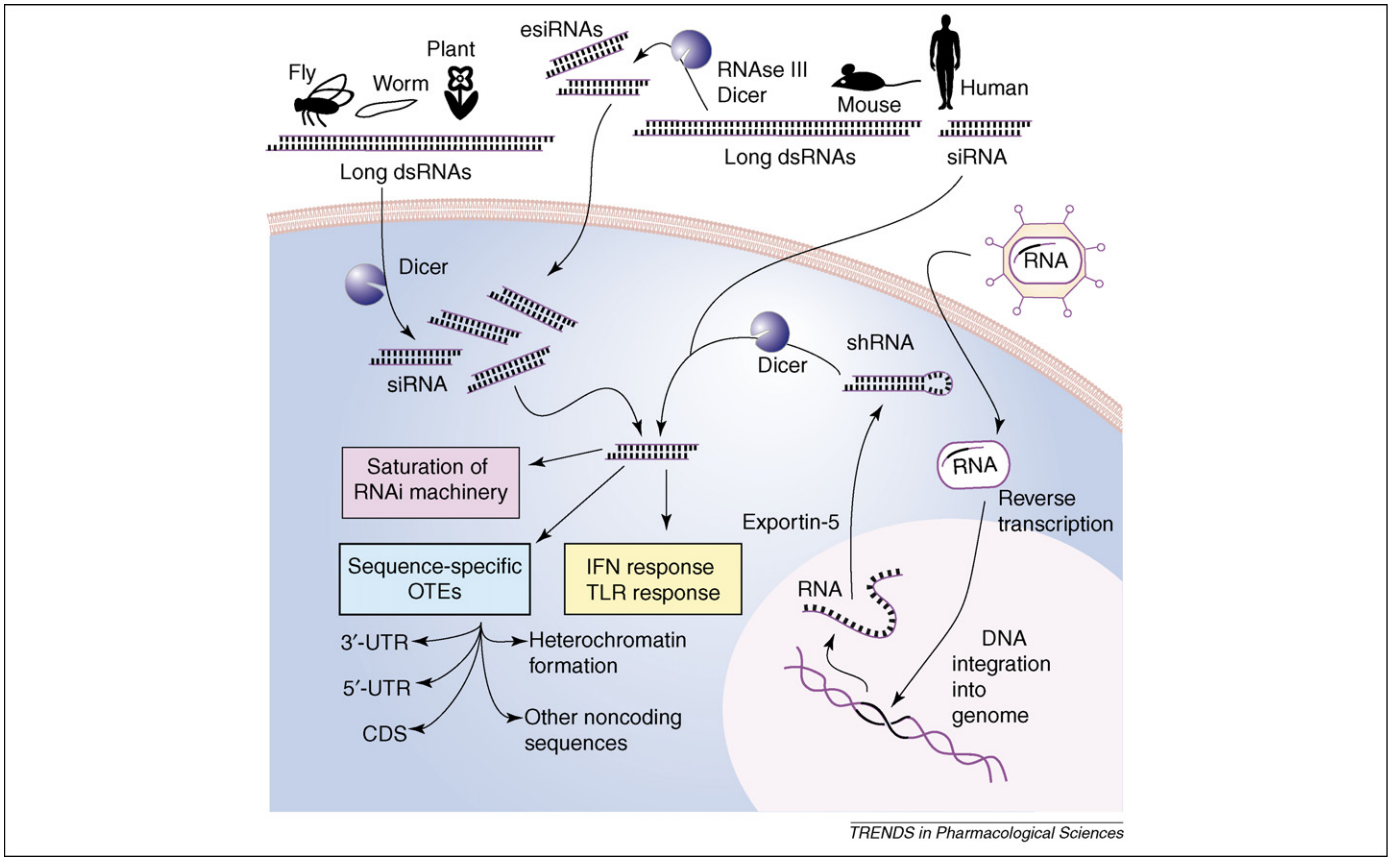

Figure 1. Origins of OTEs. Shown are the various RNAi approaches that are used in different organisms and how OTEs can arise. As the complexity and number of siRNAs used for gene silencing increase [as in the case of long dsRNAs or enzymatically prepared siRNAs (esiRNAs)] within a cell, the likelihood of inducing an OTE might also increase. Long dsRNAs are double-stranded RNA molecules that are typically several hundred base pairs in length. esiRNAs are formed from RNase-III-digested dsRNAs that are derived from DNA templates. siRNAs are ~21-nt dsRNA molecules. shRNAs are complementary stem sequences with an intervening loop region that are processed to siRNAs. Abbreviations: CDS, coding sequence; UTR, untranslated region.

cells; they looked for target genes that influence a transcriptionally responsive firefly luciferase reporter (SuperTopFlash) that contains seven consecutive copies of consensus T-cell factor (TCF)-binding sites [14]. The exclusion from their hit list of dsRNAs that were previously reported to cause growth arrest or cell death resulted in seven hits that reduced the induction of the TCF reporter >2.5-fold. Interestingly, a series of iterative mapping experiments revealed that 16–20-bp regions within the long dsRNAs were required for this effect and that each region contained distinct short homologies to *armadillo*, which is a Wg pathway component that is required for activation of the TCF reporter. Therefore, these seven potential hits were false positives. Additionally, the screen performed by Ma *et al.* identified only five of at least 17 known Wg pathway components. Together, these results indicate high rates of both false positives and false negatives for this screen. The authors raised several additional concerns about the quality of screens with first-generation long dsRNAs in *Drosophila* cells, including the reproducibility of the primary screening results with *de novo* resynthesis and retesting of the dsRNAs, cytotoxic effects caused by dsRNAs containing simple CAN (where N is any base) trinucleotide repeats and the increase in the number of OTEs with time of incubation. Overall, these findings indicate that results obtained using early versions of *Drosophila* long-dsRNA libraries should be treated with caution unless used in conjunction with redundant, non-overlapping libraries.

To gain insight into the performance characteristics of their library at a statistical level, Kulkarni *et al.* evaluated data from 30 genome-scale screens conducted at the *Drosophila* RNAi Screening Center (http://flyrnai.org/) and observed that independent, non-overlapping long dsRNAs for the same gene can have different phenotypic outcomes that are unrelated to knock-down efficiency of the target gene [15]. An algorithm search for sequences that were complementary to all possible 21-mers present in the dsRNA library showed that ~1/5 of the entire collection had one or more off-target sequences. A global analysis of multiple screens showed that genes targeted by dsRNAs with 19-nt matches elsewhere in the genome were more likely to score as hits. The authors experimentally validated the hypothesis that predicted OTEs correlate with false positives by examining the results of a screen designed to assess extracellular-signal-regulated kinase (ERK) activation. Seven candidate genes were retested using distinct dsRNAs that were computationally predicted to lack off-target sequences. Despite similar knock-down levels of the target gene being obtained using the original and the two different targeting dsRNAs, none of the seven genes could be confirmed to affect ERK activation with the newly synthesized dsRNAs. This indicates that the observed effects on ERK with first-generation dsRNAs were probably caused by OTEs. Furthermore, the authors showed that cross-hybridization of a 19-nt off-target sequence that is present in a particular protein

Box 1. Guidelines for determining RNAi specificity

For researchers planning to perform RNAi screens, suggestions of how to evaluate potential hits in a rigorous manner include [3,14,15]:

(i) confirming a hit using at least one other non-overlapping dsRNA or multiple siRNAs;
(ii) rescuing a mutant phenotype by introducing cDNA (e.g. dsRNA targeting 3′-UTRs, expression of cDNA without 3′-UTRs) or introducing silent mutations in a cDNA construct to avoid the silencing effect of dsRNA;
(iii) using chemical inhibitors to confirm phenotype (when available);
(iv) constructing dominant–negative forms of the gene of interest, if possible;
(v) using mutants (from stock collections, if available) for *in vivo* confirmation.

phosphatase 2A B′ amplicon with two predicted off-targets can lead to efficient knock-down of an unintended target. This study shows that OTEs can often be anticipated based on the dsRNA sequence but it also acknowledges that straightforward predictions about OTEs are not definitive.

Concluding remarks

The studies described demonstrate that, to minimize OTEs, long-dsRNA libraries must be optimized by including multiple distinct long dsRNAs per gene and by considering sequence-specific effects. Fortunately, second-generation libraries are being developed, and these initiatives should help to reduce false-positive rates in future screens (Box 1). The findings discussed will serve to correct some mistakes that might have arisen from long-dsRNA-based screens in the past, to enhance subsequent generations of long-dsRNA libraries and to improve the ability to employ long dsRNAs in systematic screens in the future.

Acknowledgements

We acknowledge Tom DiCesare for assistance with the figure. Our work is supported by a Human Frontiers Longterm Fellowship (to J.H.R.), a Natural Sciences and Engineering Research Council of Canada postdoctoral fellowship (to J.M.) and grants from the National Institutes of Health and the Keck Foundation (to D.M.S.). We apologize to those authors whose primary works have not been cited owing to space constraints.

References

1 Hannon, G.J. and Rossi, J.J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature* 431, 371–378
2 Cullen, B.R. (2006) Enhancing and confirming the specificity of RNAi experiments. *Nat. Methods* 3, 677–681
3 Echeverri, C.J. *et al.* (2006) Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* 3, 777–779
4 Echeverri, C.J. and Perrimon, N. (2006) High-throughput RNAi screening in cultured cells: a user's guide. *Nat. Rev. Genet.* 7, 373–384
5 Reynolds, A. *et al.* (2006) Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA* 12, 988–993
6 Sledz, C.A. *et al.* (2003) Activation of the interferon system by short-interfering RNAs. *Nat. Cell Biol.* 5, 834–839
7 Judge, A.D. *et al.* (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat. Biotechnol.* 23, 457–462
8 Hornung, V. *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
9 Bridge, A.J. *et al.* (2003) Induction of an interferon response by RNAi vectors in mammalian cells. *Nat. Genet.* 34, 263–264
10 Grimm, D. *et al.* (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441, 537–541
11 Birmingham, A. *et al.* (2006) 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Methods* 3, 199–204
12 Lin, X. *et al.* (2005) siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Res.* 33, 4527–4535
13 Jackson, A.L. *et al.* (2006) Widespread siRNA 'off-target' transcript silencing mediated by seed region sequence complementarity. *RNA* 12, 1179–1187
14 Ma, Y. *et al.* (2006) Prevalence of off-target effects in *Drosophila* RNA interference screens. *Nature* 443, 359–363
15 Kulkarni, M.M. *et al.* (2006) Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nat. Methods* 3, 833–838

0165-6147/$ – see front matter
doi:10.1016/j.tips.2007.02.009

Erratum

Erratum: Ca^{2+} channel $\alpha_2\delta$ ligands: novel modulators of neurotransmission

***Trends Pharmacol. Sci.* 28 (2007) 75–82**

In the article 'Ca^{2+} channel $\alpha_2\delta$ ligands: novel modulators of neurotransmission' by David J. Dooley, Charles P. Taylor, Sean Donevan and Douglas Feltner, which was published in the February 2007 issue of *Trends in Pharmacological Sciences*, there was an incorrect reference citation in the Figure 1 legend. Parts (a,b) were reproduced from Ref. [54], not Ref. [55]. The correct figure legend is shown below. The authors apologize to the readers for this error.

Figure 1. Coronal *in situ* mRNA hybridization and autoradiographs from similar sections of rodent brain. **(a)** *In situ* hybridization of $\alpha_2\delta$-1 mRNA in rat brain; **(b)** *in situ* hybridization of $\alpha_2\delta$-2 mRNA in rat brain; **(c)** [^{3}H]gabapentin binding (20 nM) in rat brain; **(d)** as in (c) but in the presence of unlabeled pregabalin (10 μM) to define nonspecific binding; **(e)** [^{3}H]pregabalin binding (30 nM) in wild-type mouse brain (nonspecific binding defined with gabapentin (10 μM) was not distinguishable from background (data not shown)); and **(f)** as in (e) but from a genetically altered mouse with a single amino acid mutation to $\alpha_2\delta$-1 (i.e. R217A). In (a), elevated densities of $\alpha_2\delta$-1 mRNA are localized in cell bodies of superficial neocortex, amygdala, dentate granule and pyramidal cell layers of hippocampus, and ventromedial nucleus of hypothalamus. In (b), high densities of $\alpha_2\delta$-2 mRNA are seen in cell bodies of medial habenula, reticular nucleus of thalamus, arcuate nucleus of hypothalamus, and Purkinje cell layer of cerebellum (latter not shown). In (c), high densities of [^{3}H]gabapentin binding are found in superficial neocortex, amygdala, neuropil of hippocampus, molecular layer of cerebellum (not shown) and hypothalamus. In (d), nonspecific binding is barely distinguishable from background. In (e), high densities of [^{3}H]pregabalin binding are similar to those of [^{3}H]gabapentin binding shown in (c). In (f), expression of mutant (R217A) $\alpha_2\delta$-1 protein in mice reduces binding affinity of [^{3}H]pregabalin (or [^{3}H]gabapentin) by ~90%; note decreased intensity of binding in neocortex, amygdala, hippocampus and ventromedial nucleus of hypothalamus. Abbreviations: Amy, amygdala; ARH, arcuate nucleus of hypothalamus; CA1, hippocampal subfield CA1; CA3, hippocampal subfield CA3; CP, caudate putamen; DG, dentate gyrus; MH, medial habenula; RT, reticular nucleus of thalamus; Th, thalamus; VMH, ventromedial nucleus of hypothalamus. (a,b) Reproduced, with permission, from Ref. [54]; (c,d) used with permission of Y. Dumont and R. Quirion (unpublished data); (e,f) reproduced, with permission, from Ref. [19].

☆ DOI of original article: 10.1016/j.tips.2006.12.006.
Available online 6 March 2007.

0165-6147/$ – see front matter
doi:10.1016/j.tips.2007.02.001