OPINION

Unravelling the general properties of siRNAs: strength in numbers and lessons from the past

Jonathan Hall

The optimal use of small interfering RNAs (siRNAs) requires an understanding of their general properties, and particularly their selectivity and potency. However, it is often difficult to distinguish the properties of individual siRNAs from their general properties as a class of molecule. The analysis of large sets of siRNAs is one solution. Moreover, years of research into the general properties of antisense oligonucleotides have provided some valuable pointers for designing experiments to unravel the general properties of this new generation of gene-silencing oligonucleotides.

RNA INTERFERENCE (RNAi) represents one of the most powerful biological tools ever to be introduced. It provides a simple, rapid, inexpensive, selective method of gene inhibition with high success rate. Gene-specific RNAi screens in cells or model organisms generate data that link a specific gene to a given biological process. Moreover, genome-wide screens that use a library of individual oligoribonucleotides to knock down each gene return massive amounts of biological information. The results from the first genome-wide libraries of RNAi reagents — dsRNAs in the model organism *Caenorhabditis elegans*[1] and in *Drosophila melanogaster* cell culture[2] — are now available.

Each dsRNA that is specific for a given gene in model-organism RNAi libraries is typically hundreds of nucleotides in length. Each of these dsRNAs is enzymatically processed into a large population of SMALL INTERFERING RNAs (siRNAs), which downregulate levels of the target mRNA. Intuitively, this large number of individual siRNAs should allow each dsRNA in a library to knock down the homologous gene with high sequence specificity and high potency (functionality). However, in most mammalian systems, long dsRNAs seem to induce a toxic INTERFERON RESPONSE. Nonetheless, siRNAs themselves do not elicit this response, presumably because they are too short (see REF. 3).

Within months of the first reports that siRNAs could be used to inhibit gene expression in mammalian cells, a small series of articles and commentaries had proclaimed that this approach was superior to alternative ANTISENSE techniques[4,5]. At miniscule concentrations, siRNAs were said to provide an almost-perfect success rate and an extraordinary specificity, therefore opening the doors to new therapies and to new techniques for genome-wide functional analysis[6].

However, as the application of siRNAs started to become more widespread, reports of shortcomings of the technique increased: these mainly involved associated toxicity, poor levels of inhibition and numerous 'off-target' effects. Many of these reports were contradictory, with conclusions from one study being strongly contested by the next. Here, I argue that these perceived disparities result from the misconception that the behaviour of one oligonucleotide under a certain set of experimental conditions describes the behaviour of this class of molecule under a broad range of conditions. Oligonucleotides do indeed share common properties as a class, but each sequence will also have individual properties, such as potency and stability, that are determined by the specific nucleotide sequence. This has not presented a hurdle for the routine laboratory use of siRNAs as a biological tool, as reagents are usually characterized before initiating extensive experimentation. Reagents that perform unsatisfactorily are simply replaced.

However, genome-wide libraries of human RNAi reagents for high-throughput gene screening will comprise a single (or a small number of) siRNA(s) per gene, applied to cells at relatively high concentrations to ensure potent effects (FIG. 1). So, for each targeted gene, a single siRNA needs to be selected from thousands of candidate siRNAs. Therefore, each siRNA should be potent and specific. The size of such libraries prohibits the characterization of individual siRNA members before experimentation, and, therefore, the design of a genome set needs to take into account the general properties of siRNA oligoribonucleotides as a class. Two such properties are crucial: selectivity at the single-nucleotide level and potency.

The first large-scale gene-knockdown experiments with short hairpin RNAs (shRNAs) in mammalian cells confirm that gene-screening techniques can yield insight into mammalian gene function: in 1 case, 4,873 genes were screened in retroviruses to identify new genes that interfere in the proteasome pathway[7], and in the other, 7,914 genes were screened using a library of shRNA pools to uncover 5 new modulators in the p53 pathway[8].

Here, I argue that to best learn about the general properties of siRNAs that induce RNAi, several experiments need to be performed from which large homogeneous data sets can be generated. From the analysis of such data sets, patterns of general behaviour emerge that are not apparent from the

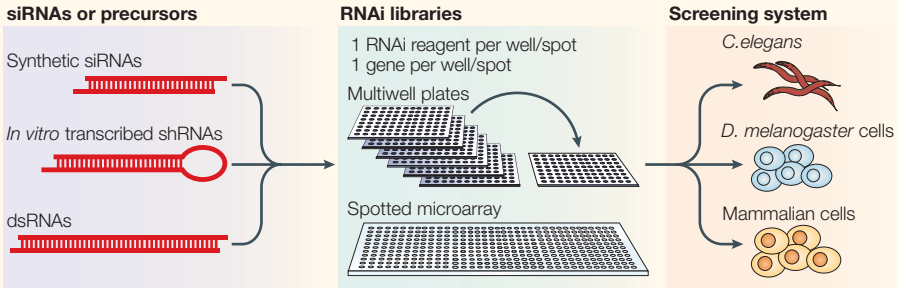


Figure 1 | **Genome-wide screens.** So far, large-scale RNA interference (RNAi) screens have been performed in *Caenorhabditis elegans*, *Drosophila melanogaster* and mammalian cells. Libraries of RNAi reagents have been created from various sources: small interfering RNAs (siRNAs) are synthesized chemically or prepared enzymatically from primers with T7-RNA polymerase[54]; SHORT HAIRPIN RNAs (shRNAs) are prepared by transcription from Pol II/III promoters and are converted to siRNAs intracellularly by the Dicer enzyme[55,56]; long dsRNAs are prepared by PCR and are processed into siRNAs by Dicer. For each targeted gene, 1–5 siRNAs, or siRNA equivalents, are stored in individual wells of microtitre plates. Reagents are transferred using robotics into microtitre plates that contain living worms or cells, or onto the surface of arrays before the addition of cells.

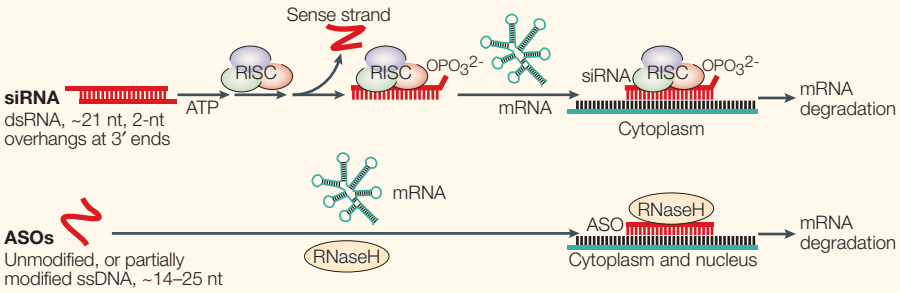


Figure 2 | **Gene inhibition by antisense oligonucleotides and small interfering RNAs.** Small interfering RNA (siRNA)-induced degradation of mRNA in the cytoplasm comprises four principal steps[57]: ATP-independent incorporation of siRNAs into the multi-subunit RNA-INDUCED SILENCING COMPLEX (RISC; possibly in association with the ribosome[58,59]); ATP-dependent unwinding of the siRNA duplex; ATP-independent binding to mRNA; and irreversible cleavage by RNase-III-type activity as the first step of mRNA degradation. Antisense oligonucleotides (ASOs) function as gene inhibitors by one of several mechanisms, depending on the chemical composition of the oligonucleotide. The most commonly used examples are partially modified ASOs, which are delivered to cells in single-stranded form and induce RNase H to bind to the mRNA–oligonucleotide complex in the cytoplasm and nucleus. RNase H cleaves the mRNA progressively. Both processes occur with catalytic amounts of oligonucleotides, although under conditions of cellular transfections, oligonucleotides are often present in excess quantities over the target mRNA. The principal difference between the two mechanisms is the nature and role of the enzymes. However, the interaction of both these types of oligonucleotide with the mRNA follows the rules of Watson–Crick recognition.

behaviour of small numbers of siRNAs in isolated experiments. Moreover, I argue that much can be learned from the antisense field and previous experiments that attempted to address the same questions for antisense oligonucleotides (ASOs; FIG. 2). I review recent work on mismatch selectivity, microRNA (miRNA)-based inhibition, off-target effects and the mechanism of action of the RNA-induced silencing complex (RISC) to illustrate examples of both the individual and the general properties of siRNAs. I also highlight the value of previous antisense research to this area and summarize what we currently know about the general properties of siRNAs (for a recent comprehensive review of advances in RNAi technology, see REF. 9).

Specificity

Single-nucleotide selectivity. Probably the best guide that we have to what determines the nucleotide specificity of siRNAs is the body of work that is already available on the specificity of ASOs (BOX 1). In general, these studies show that oligonucleotides are not 'specific', but rather are 'selective'. So, although a single mismatch in an ASO could theoretically lead to a 500-fold decrease in target affinity[10], in the whole-cell environment, several compensatory factors, such as higher concentrations of highly homologous mRNAs, can result in unintended inhibition of non-target genes (see examples in REFS 11,12). Such off-target effects strongly depend on oligonucleotide length, chemical modification, target accessibility, percentage homology and concentrations of reagent and mRNA[10]. Excellent single-nucleotide selectivity using ASOs is possible, but it requires careful experimental optimization[13].

Similar to an ASO, a siRNA or a shRNA should have a maximum number of mismatches to all other potential binding sites in the transcriptome to ensure the highest possible sequence selectivity. Therefore, the design of a selective inhibitor requires an understanding of the general behaviour of siRNAs towards binding sites with only a few mismatches.

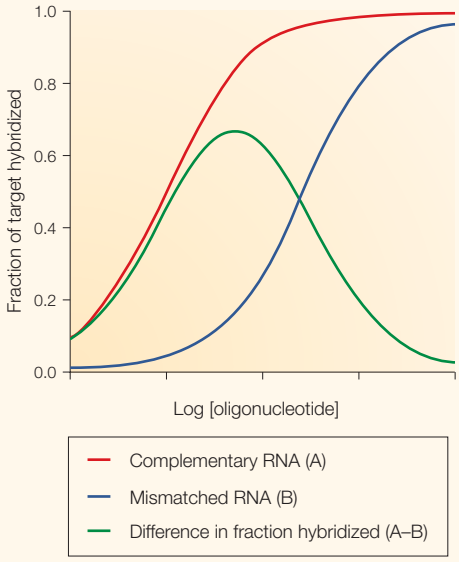
Early publications on siRNAs optimistically asserted that a single-nucleotide mismatch was sufficient to render a siRNA duplex inactive in mammalian cells[14,15]. Recent publications have now demonstrated experimentally that such single-nucleotide selectivity is attainable. For example, a mutant disease-causing allele of the spinocerebellar ataxia type 3 (Machado–Joseph disease; *MJD*) gene was selectively silenced in cell culture[16] by incorporating a mismatch at position 10 of the target site. Selective silencing of alleles of the muscle acetylcholine receptor (*ACHR*) has also been reported[17], with a significant but incomplete degree of selective inhibition of the mutated gene observed at the protein level (83% inhibition compared with 37%) after incorporation of a mismatch at position 9. So, as for ASOs, siRNA selectivity at the single-nucleotide level is possible, but it requires extensive optimization.

A broader investigation of single-nucleotide selectivity using 10–20 siRNAs supported these observations[18]. In this case, siRNAs were designed with single base-pair mismatches to the mRNA target site at 6 different 5′ locations and 2 different 3′ locations. Only the 3′ mismatches showed significantly less inhibition compared with the perfectly matched siRNA. Chiu and Rana[19] also found 2 examples of mismatches at the 3′ end providing superior selectivity. By contrast, a systematic analysis of shRNAs with single base-pair mismatches to the 21 positions in a target site in the mRNA of the human immunodeficiency virus 1 *Gag* gene showed that mismatches at the central target positions 9–11 and at positions 4 and 16–18 were the most discriminatory[20].

Collectively, these efforts have not fully explained the general behaviour of siRNAs with respect to single-nucleotide specificity. They demonstrate convincingly that fully complementary reagents are the most inhibitory and that mismatches at terminal locations provide only minor discrimination. However, the data sets are too small to conclude whether, in general, there is a minimum number of mismatches required to ensure specificity for a target gene under a broad range of conditions or if different regions of the oligoribonucleoides (5′, 3′ regions, target cleavage site) contribute differently to selectivity. Nevertheless, bioinformatic analyses that are aimed at identifying the mRNA targets of miRNAs do indicate that some regions in such oligoribonucleotides might be more important than others[21], and new experimental evidence supports this[22]. Regardless, the few studies discussed above indicate that if RNAi libraries are designed for high-throughput screening, reagents probably need to have multiple mismatches to all other potential binding sites in the transcriptome to minimize off-target effects.

siRNAs and miRNAs. An added complication to siRNA design stems from recent suggestions[23] that, similar to miRNAs, siRNAs might be able to bind to partially homologous sites that contain insertions and deletions and still inhibit gene expression. miRNAs are a second class of small regulatory RNAs that inhibit gene expression by non-degradative translational attenuation. Scacheri *et al.*[23] proposed that siRNAs might act through a miRNA-like mechanism to explain their finding that a subset of the siRNAs designed to knockdown the *MEN1* (multiple endocrine neoplasia I) gene significantly upregulated *p21* (cyclin-dependent kinase inhibitor 1A; *CDKN1A*) and *p53* (tumour protein 53; *TP53*) genes. However, bioinformatic analyses did not reveal any probable miRNA targets, so the real origin(s) of these effects remain unknown.

Box 1 | **Selectivity of oligonucleotides**



The figure shows the theoretical *in vitro* hybridization profiles of an oligonucleotide that is present in excess to a fully complementary RNA (A) and to a partially complementary RNA (B) that bears a mismatch. In each case, the fraction of bound RNA varies with the affinity of the oligonucleotide for the target, and the concentration. The selectivity of the oligonucleotide for RNA (A) over (B) is therefore determined by concentration and affinity and, for any two given RNAs, is possible within a concentration range that is defined by the line (C). As the horizontal distance between the curves (A) and (B) increases, the concentration range for which good selectivity can be obtained is widened. So, to maximize the selectivity of an oligonucleotide for its complementary RNA in the presence of highly homologous RNAs, an antisense sequence should be designed to have the greatest difference in affinity (that is, the highest melting temperature difference; ΔTm) with the next closest predicted binding homologue. So, the oligonucleotide should have the maximum affinity for the target and a maximum number of mismatches to any other closely related homologues, such that curve (A) can be shifted to the left and curve (B) to the right.

In a gene-knockdown experiment that uses oligonucleotides that induce mRNA cleavage, specific target cleavage rather than specific target binding, is the goal. Factors that affect target cleavage are numerous and complicated. These factors include the concentrations, and secondary and tertiary structures of target and non-target mRNAs, as well as sequence motifs. Consequently, theoretical selectivity that is based on the fraction of the target that is bound might not reflect selectivity of the fraction of the target that is actually cleaved.

Nevertheless, for antisense oligonucleotides (ASOs) hybridization thermodynamics is a crucial factor for efficient, selective gene inhibition and explains cases in which good selectivity has been observed[49]. In recent years, several structural modifications to ASOs that provided improved mismatch specificity were introduced[50], and, in contrast to small interfering RNAs (siRNAs), the length of ASOs can be varied to further improve selectivity[51].

The situation for siRNAs is different. The RNA interference (RNAi) mechanism involves multiple discrete enzymatic steps before hybridization of siRNA with target mRNA. Moreover, RNA-induced silencing complex (RISC) elements that are responsible for mRNA cleavage might even aid the hybridization process. By contrast, ASOs probably recruit RNase H only after hybridization. Furthermore, emerging evidence from microRNAs (miRNAs)[22] indicates that different regions of a RISC-loaded guide strand have different roles in RNAi. Importantly, the 5′ half of the guide is more involved in target recognition, and therefore is the best location in which to incorporate mismatches to other highly homologous mRNA binding sites. Although these factors taken together imply that the above model might be less applicable to siRNAs than ASOs, Pancoska *et al.*[52] have provided experimental evidence that hybridization thermodynamics between the siRNA guide strand and the mRNA is an important factor for efficacious gene inhibition and that the excellent selectivity that was observed in the experiments of Semizarov *et al.* included a sequence design that incorporated hybridization thermodynamics. In conclusion, the mechanism of action of the individual oligonucleotides should define the best approach for investigators to use for target-specific inhibition. Figure modified with permission from REF. 53 © (1993) CRC Press.

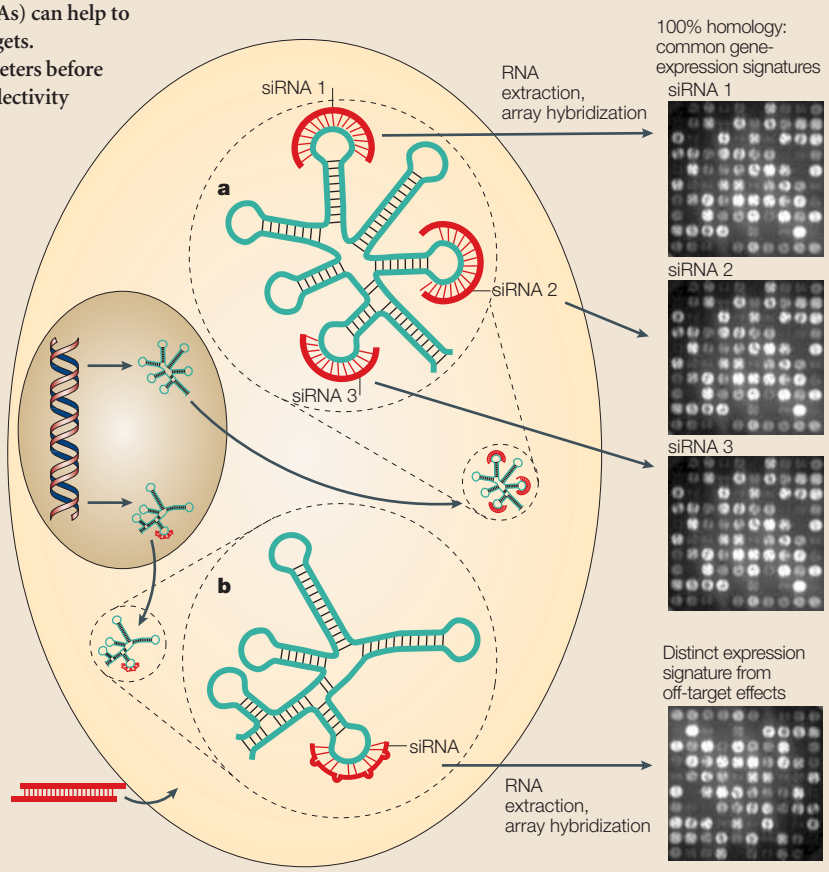
A handful of papers have described experiments that were performed to specifically investigate gene inhibition by small RNAs that act at partially complementary mRNA sites. In one study, co-transfection of a plasmid that codes for *miR-30*, a naturally-occuring miRNA that is expressed in mice and humans, and a luciferase reporter gene that carries 4 identical, consecutive partially homologous target sites in the 3′-untranslated region (UTR), resulted in expression of *miR-30* and inhibition of luciferase activity without mRNA degradation[24]. If the bulged target sites were exchanged for a single fully complementary site, luciferase activity was inhibited through mRNA degradation. These experiments showed that miRNAs can function as siRNAs. But can siRNAs function as miRNAs? A second publication from the same group[25] answered this question. Transfection of a siRNA homologous to the *Drosophila melanogaster Nxt* gene inhibited translation of luciferase that has eight identical repeated miRNA-like target sites, without affecting mRNA levels.

Independent confirmation that siRNAs can downregulate genes by interacting with bulged target mRNAs came from another group[26]. Four consecutive target sites that were predicted to form bulged duplexes with a *CXCR4* (chemokine (C-X-C motif) receptor 4) siRNA were introduced into the 3′ UTR of a luciferase gene. Again, good inhibition of luciferase activity was achieved without any observed lowering of mRNA levels. Additional experiments to define the rules of the interactions led the authors to suggest that the sequence of the bulge is not a principal determinant, that a minimum amount of free energy for binding the first 8 nucleotides of the miRNA 5′ region is necessary for translational inhibition and that complementarity in the 3′ end of the miRNA is less crucial than at the 5′ end[22].

Together, these three studies show that siRNAs can downregulate genes in a miRNA-like fashion. However, multiple repeat copies of the target site were required to see efficient translational inhibition. This in itself should ensure that such interactions are unlikely to pose particular problems with respect to siRNA design. However, Saxena *et al.*[27] reported that siRNAs with partial complementarity to a single site in the coding regions of *p21* and *geminin* (*GMNN*) caused translational inhibition. In this case, similar levels of inhibition were observed at complementary and bulged mRNA target sites, and inhibition occurred through interaction at a single binding site, located in the coding region. If more examples that are similar to this emerge, then this non-complementary interaction will indeed pose problems for the design of selective siRNA reagents until the rules that govern the interaction are better understood.

Genome-wide studies. DNA arrays provide a snapshot of the mRNA levels of all cellular genes represented by oligonucleotide probes that are present on the array. Therefore, they provide a comprehensive means to query the selectivity of an individual oligoribonucleotide in the cell, albeit with the limitation that the technique only records changes in mRNA levels: that is, translational inhibition is not directly observed. Such studies can also

Box 2 | **Investigating the selectivity of small interfering RNAs using gene-expression profiling**



Careful sequence design of small interfering RNAs (siRNAs) can help to ensure that reagents are selective for intended mRNA targets.

For example, Semizarov *et al.*[30] optimized several parameters before carrying out array hybridizations to achieve maximum selectivity of siRNAs. Three target genes were selected. Multiple characterized siRNAs per mRNA were used: various 19-nucleotide sequences against the target genes were ordered by decreasing homology against the next best hit within the transcriptome, and only those with fewer than 15 nucleotides to the next best hit were selected. Those siRNAs with the highest predicted difference in calculated affinity between the match target and the next most probable predicted mismatched mRNA target were then chosen. The authors applied the same selection criteria to the mismatched negative controls. Transfection of siRNAs at 100 nM usually led to the induction of selected apoptosis and stress-response genes. However, a short-dose response and time-course study showed that transfection of the siRNAs at 20 nM was equally inhibitory against the intended target, whereas most of the previously observable nonspecific effects disappeared, and, therefore, siRNAs were used at this concentration. Transfection of multiple siRNAs that targeted any one of the given genes, as depicted in (a), then resulted in highly correlated gene-expression signatures.

Under non-optimized experimental conditions, siRNAs, like antisense oligonucleotides, bind and inhibit the expression of genes that contain partially complementary target-binding sites (b). A gene-expression profile that uses DNA microarrays yields a distinct expression signature that is associated with the 'off-target' effects of the siRNA (b)[29].

provide gene-expression signatures for groups of siRNAs that are directed to distinct sites of a given gene. In these cases, modified expression of a common set of genes from the relevant pathways is observed in addition to target-gene inhibition.

Four publications describe microarray experiments performed to address selectivity of siRNAs. In the first, microarray analyses of transfections of *GFP* siRNAs into a *GFP*-expressing cell line showed that out of 36,000 genes, only *GFP* was consistently knocked down, which indicated that siRNAs were highly specific[28]. The contradictory results and conclusions of two more papers in 2003 added fierce controversy to the debate on siRNA specificity. Genes with well-characterized biology were targeted in both studies (BOX 2). In one, 24 siRNAs, designed by standard selection rules[14] and with fewer than 18 nucleotides of homology against any other theoretical target sites, were used to silence two genes, *IGF1R* (insulin-like growth factor 1 receptor) and *p38* (REF. 29). Each of the siRNAs produced distinct gene-expression signatures and none of the down-regulated genes was a recognized member of the targeted pathways. The authors suggested that cross-hybridization with partially homologous sites (with as few as 11 shared nucleotides) led to these off-target effects, and that a siRNA sequence that is 'absolutely' specific would therefore be difficult to select.

The microarray experiments of Semizarov *et al.*[30] began after a careful analysis of reagents and experimental conditions. Five siRNAs were used for each of the three target genes. *RB1* siRNAs modified the expression of 2,475 genes in these experiments. All five *RB1* siRNAs had similar gene-expression signatures, which included a modified expression of the target cell cycle and DNA biosynthesis genes. Similarly high correlations were obtained for *AKT1* and *PLK1* siRNAs. siRNAs against different targets showed little overlap in expression signature, confirming that effects were due to specific target downregulation. So, in contrast to the conclusions from the *IGF1R* and *p38* study, these data indicate that siRNAs can be highly specific in mammalian cells with careful siRNA selection, meticulous experimental design and the inclusion of negative controls.

The last publication of the four describes microarray experiments in which luciferase-expressing cells were over-dosed with an anti-luciferase siRNA[31]. Large numbers of non-target genes were up- and downregulated, and 12 of these were selected for more detailed study. A second unrelated siRNA regulated the 12 genes in a similar fashion, and a subset of these was similarly affected by treatment of cells with interferon or with long dsRNAs. As the 12 genes responded identically to 2 unrelated siRNAs, it might be tempting to assume that they are regulated by siRNAs in general. However, these effects might simply have been caused by the chance presence of a single motif in both siRNAs, just as a CpG motif in ASOs or a terminal triphosphate on a siRNA can cause distinct nonspecific effects[11,12,32]. More experiments using more siRNAs together with a set of oligonucleotide controls, such as plasmid DNA or ASOs, are needed to test this hypothesis.

Genome-wide microarray experiments are too sophisticated and expensive to consider as a routine method to assess the selectivity of any given siRNA and to allow researchers to

be sure that phenotypic effects are due to the inhibition of the intended target. There is sufficient choice of other types of control available to do this: multiple siRNAs that are targeted to distinct non-overlapping sites of the same target, cDNA-rescue constructs that are unaffected by the siRNA in question and unrelated siRNA-negative controls. However, microarray studies might be required for siRNAs designed as therapeutic agents to help to assess potential toxicity[33].

Potency

The functionality of si/shRNAs has been one of the most contentious aspects of the field. Oligonucleotide potency, similar to selectivity, is a subjective quality that depends on numerous experimental parameters. So, references to potency make most sense if used in a relative context in which oligoribonucleotides are normalized to a reference, or to each other. Early controversy in the field centred on claims of an almost-perfect success rate with siRNAs. Harborth *et al.*[14] showed that the first siRNAs selected downregulated 14 out of 16 endogenously-expressed genes, including Lamin A/C (*Lmna*) (see also REF. 34). Not everyone experienced this success. For example, only 1 out of 5 *Lmna* siRNAs were effective in a subsequent study[35]. Similarly, Holen *et al.*[36], in the first systematic evaluation of the efficacy of multiple siRNAs against the same target mRNA, found that siRNAs showed varying levels of functionality. Another early larger-scale study compared the potency of siRNAs to ASOs in 2 genes (*CD54* (*ICAM1*) and *PTEN*)[37]. Only one-third of the ~80 siRNAs tested were classified as 'active': half as many as the ASOs.

Today, with hundreds of publications that describe the use of siRNAs, it is apparent that many are indeed non-functional, whatever the experimental conditions. This constitutes a serious hurdle for the construction of genome-wide collections as investigators are obliged to include multiple si/shRNAs in a library to ensure efficient knockdown of each targeted gene, adding significantly to the expense of reagents, screening and data analysis[7,8].

Two back-to-back papers that describe how the RISC selects the antisense guide strand have contributed significantly to the understanding of features that confer potency to a siRNA. The affinity of oligonucleotide–mRNA-duplex formation has long been known to be essential for an efficient antisense-based inhibition, and early RNAi work indicated that duplex unwinding is a crucial processing step of both dsRNA and pre-miRNAs. Khvorova *et al.*[38] examined the theoretical thermodynamics profiles of 272 miRNAs and their predicted duplexes with mRNAs. They discovered that the 5′ terminus of the guide strand of the miRNA consistently showed a weaker predicted binding affinity than the corresponding 3′ end. The related enzymatic processing of miRNAs and siRNAs led them to predict that the selection criteria for the siRNA guide strand might be similar for the stabilization of miRNAs by RISC. Visual analysis of thermodynamic properties of 37 active siRNAs showed that the 5′ end of the antisense strand of the potent siRNAs also consistently showed a weaker predicted duplex-binding potential than that of the sense strand. This observation was confirmed with 3 sets of experiments that involved more than 200 siRNAs: the strongest inhibitors all carried the predicted profile. So, the analysis of large data sets, first from miRNAs and then from siRNAs, revealed a truly general property of potent siRNAs. In future, this paper will probably significantly influence the design of siRNA reagents.

Another study that was published at the same time provided supporting evidence for the results reported by Khvorova *et al.* and an explanation as to why this local duplex affinity is a key determinant of RISC selection[39]. Terminal nucleotides were mutated on both strands of a siRNA to change the binding affinity of the oligoribonucleotides: the siRNA strand with the less stable 5′ end is preferentially incorporated into the RISC complex.

A more recent study of 180 randomly selected siRNAs from regions poor in G+C content, targeting 2 genes, has built on the conclusions of these 2 papers[40]. Once again, the large data sets revealed generic traits, including positive and negative determinants. For example, a 2-nucleotide shift in the targeted region can cause a large change in functionality, and approximately 78% of the sequences induced more than 50% silencing. In the sense strand of active siRNAs, an A+U-rich region is observed in the 3′ part: A is often found at positions 3 and 19, and U at position 10. These features were combined and incorporated into an algorithm-based design tool that was aimed at improving the selection of potent reagents.

Three other reports have used large sets of reagents to investigate the general properties of potent siRNAs. A study of 62 siRNAs highlighted several common features of the antisense strand of active siRNA sequences: A or U at the 5′ terminus, G or C at the 3′ end and an absence of G+C stretches of more than 9 nucleotides[41]. Similarly, a study of approximately 150 siRNAs targeted at 22 genes of the PI3K (phosphatidylinositol 3-kinase) pathway revealed a preference for U at the 5′ terminus (position 1) and G or C at position 9 of the antisense strand[42]. In the last of these recent studies, Amarzguioui and Prydz[43] evaluated 46 siRNAs and found that A at position 6 can be helpful, in addition to the A+U-rich region at the guide strand 5′ terminus.

This small group of publications demonstrates how larger-than-average data sets have been used to identify general features of potent siRNAs, and how they have been included into effective selection algorithms. Investigators agree on the importance of an A+U-rich region at the 5′ end of the guide strand, including the terminal position, but, at present, there seems to be little consensus for other positions. It is probable that more consensus sequence motifs will emerge, possibly more complex than at the single-nucleotide level, as data sets continue to grow in size and more sophisticated analysis methods are applied to the data. Ultimately, the use of neural network programmes will be used with such data sets, as was recently described for ASOs[44], and the outcome will be a tool that predicts potent siRNAs with high accuracy.

Glossary

ANTISENSE
DNA or RNA that is manipulated in a laboratory to be complementary to a target mRNA. Antisense techniques are used to inhibit the expression of genes in a sequence-specific fashion.

INTERFERON RESPONSE
A primitive antiviral mechanism that triggers sequence-nonspecific degradation of mRNA and downregulation of cellular protein synthesis.

microRNA
(miRNA). Small regulatory, antisense RNAs (21–25 nucleotides long) that repress the translation of homologous target RNA.

RNA-INDUCED SILENCING COMPLEX
(RISC). A multi-component, ribonucleoprotein complex that cleaves specific mRNAs that are targeted for degradation by homologous dsRNAs during the process of RNA interference.

RNA INTERFERENCE
(RNAi). A process by which dsRNA specifically silences the expression of homologous genes.

SHORT HAIRPIN RNAs
(shRNAs). Small RNAs that form hairpins that can induce sequence-specific silencing in mammalian cells through RNA interference, both when produced exogenously and transfected into the cell, and when expressed endogenously.

SMALL INTERFERING RNAs
(siRNAs). Small antisense RNAs (20–25 nucleotides long) that are generated from specific dsRNAs that trigger RNA interference. They serve as guides for the cleavage of homologous mRNA in the RNA-induced silencing complex (RISC).

Investigators cannot agree on the importance of mRNA secondary and tertiary structure to siRNA potency[40,45]. The importance of secondary structure was elegantly addressed previously for ASOs by cloning an invariant target site into mRNA constructs of well-defined secondary structures, and monitoring the activity of a constant ASO as a function of changing structure: stable RNA structure led to attenuated inhibition[46]. In the same system, active siRNAs behaved similarly, and, therefore, it is probable that target structure also affects siRNAs[37]. A much larger study that used approximately 50 different target constructs confirmed this finding[47]. These results imply that, although including particular sequence motifs when designing siRNAs can optimize their potency, the activity of siRNAs is probably never completely independent from the effects of mRNA local structure.

Conclusions

Investigators have struggled to explain two general properties of siRNAs: potency and selectivity. The more we know about these properties, the better will be the performance of what is already a highly effective and widely used tool. Moreover, a deep understanding of these properties is essential for the design of si/shRNAs for genome-wide screening (FIG. 1). There are now hundreds of publications that describe the use of siRNAs in biology, and although several of these purport to portray the general properties of siRNAs, those that use large data sets can reliably reveal the general properties of these reagents as a class, as opposed to the properties of individual sequences. Furthermore, in many cases, experimental design has been most effective if we have taken heed of lessons from the past[48].

Jonathan Hall is at the Novartis Pharma AG, Department of Functional Genomics, WSJ88.431, CH-4002 Basel, Switzerland.
e-mail: jonathan.hall@pharma.novartis.com

doi:10.1038/nrg1382

1. Ashrafi, K. *et al.* Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* **421**, 268–272 (2003).
2. Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–836 (2004).
3. Caplen N. J., Parrish S., Imani F., Fire A. & Morgan R. A. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl Acad. Sci. USA* **98**, 9742–9747 (2001).
4. Bass, B. L. RNA interference: the short answer. *Nature* **411**, 428–429 (2001).
5. Howard, K. Unlocking the money-making potential of RNAi. *Nature Biotechnol.*, **21**, 1441–1446 (2003).
6. Frankish, H. Consortium uses RNAi to uncover genes' function. *Lancet* **361**, 584 (2003).
7. Paddison, P. *et al.* A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427–431 (2004).
8. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
9. Mittal, V. Improving the efficiency of RNA interference in mammals. *Nature Rev. Genet.* **5**, 355–365 (2004).
10. Freier, S. M. *et al.* in *Gene Regulation by Antisense Nucleic Acids* (Raven Press Series on Molecular and Cellular Biology, Vol. 1) (eds Ivant, J. & Erickson, R.) 95–107 (Raven, New York, 1992).
11. Tidd, D. M. & Giles, R. V. In *Pharmaceutical Aspects of Oligonucleotides* (eds Couvreur, P. & Malvy, C.) Part I, 3–31 (Taylor and Francis, London, 2000).
12. Stein, C. A. Is irrelevant cleavage the price of antisense efficacy? *Pharmacol. Ther.* **85**, 231–236 (2000).
13. Giles, R. V., Ruddell, C. J., Spiller, D. G., Green, J. A. & Tidd, D. M. Single base discrimination for ribonuclease H-dependent antisense effects within intact human leukemia cells. *Nucleic Acids Res.* **23**, 954–961 (1995).
14. Harborth, J., Elbashir, S. M., Bechert, K., Tuschl, T. & Weber, K. Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J. Cell Sci.* **114**, 4557–4565 (2001).
15. Elbashir, S. M., Harborth, J., Weber, K. & Tuschl, T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**, 199–213 (2002).
16. Miller, V. *et al.* Allele-specific silencing of dominant disease genes. *Proc. Natl Acad. Sci. USA* **100**, 7195–7200 (2003).
17. Abdelgany, A., Wood, M. & Beeson, D. Allele-specific silencing of a pathogenic mutant acetylcholine receptor subunit by RNA interference. *Hum. Mol. Genet.* **12**, 2637–2644 (2003).
18. Amarzguioui, M., Holen, T., Babaie, E. & Prydz, H. Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res.* **31**, 589–595 (2003).
19. Chiu, Y.-L. & Rana, T. M. siRNA function in RNAi: a chemical modification analysis. *RNA* **9**, 1034–1048 (2003).
20. Pusch, O. *et al.* Nucleotide sequence homology requirements of HIV-1-specific short hairpin RNA. *Nucleic Acids Res.* **31**, 6444–6449 (2003).
21. Lewis, B. P., Shih, I.-h., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
22. Doench, J. G. & Sharp, P. A. Specificity of microRNA target selection in translational repression. *Genes Dev.* **18**, 504–511 (2004).
23. Scacheri, P. C. *et al.* Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 1892–1897 (2004).
24. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed microRNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cell* **9**, 1327–1331 (2002).
25. Zeng, Y., Yi, R. & Cullen, B. R. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl Acad. Sci. USA* **100**, 9779–9784 (2003).
26. Doench, J. G., Petersen, C. P. & Sharp, P. A. siRNAs can function as miRNAs. *Genes Dev.* **17**, 438–442 (2003).
27. Saxena, S., Jonsson, Z. O. & Dutta, A. Small RNAs with imperfect match to endogenous mRNA repress translation: implications for off-target activity of small inhibitory RNA in mammalian cells. *J. Biol. Chem.* **278**, 44312–44319 (2003).
28. Chi, J.-T. *et al.* Genome-wide view of gene silencing by small interfering RNAs. *Proc. Natl Acad. Sci. USA* **100**, 6343–6346 (2003).
29. Jackson, A. L. *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnol.* **21**, 635–637 (2003).
30. Semizarov, D. *et al.* Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl Acad. Sci. USA* **100**, 6347–6352 (2003).
31. Persengiev, S. P., Zhu, X. & Green, M. R. Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). *RNA* **10**, 12–18 (2004).
32. Kim, D.-H. *et al.* Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nature Biotechnol.* **22**, 321–325 (2004).
33. Ganju, P. & Hall, J. Potential applications of siRNA for pain therapy. *Expert Opin. Biol. Ther.* **4**, 531–542 (2004).
34. Harborth, J. *et al.* Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev.* **13**, 83–105 (2003).
35. Kumar, R., Conklin, D. S. & Mittal, V. High-throughput selection of effective RNAi probes for gene silencing. *Genome Res.* **13**, 2333–2340 (2003).
36. Holen, T., Amarzguioui, M., Wiiger, M. T., Babaie, E. & Prydz, H. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res.* **30**, 1757–1766 (2002).
37. Vickers, T. A. *et al.* Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *J. Biol. Chem.* **278**, 7108–7118 (2003).
38. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–216 (2003).
39. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208 (2003).
40. Reynolds, A. *et al.* Rational siRNA design for RNA interference. *Nature Biotechnol.* **22**, 326–330 (2004).
41. Ui-Tei, K. *et al.* Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference *Nucleic Acids Res.* **32**, 936–948 (2004).
42. Hsieh, A. C. *et al.* A library of siRNA duplexes targeting the phosphoinositide 3-kinase pathway: determinants of gene silencing for use in cell-based screens. *Nucleic Acids Res.* **32**, 893–901 (2004).
43. Amarzguioui, M. & Prydz, H. An algorithm for selection of functional siRNA sequences. *Biochem. Biophys. Res. Commun.* **316**, 1050–1058 (2004).
44. Giddings, M. C. *et al.* Artificial neural network prediction of antisense oligodeoxynucleotide activity. *Nucleic Acids Res.* **30**, 4295–4304 (2002).
45. Kretschmer-Kazemi, F. R. & Sczakiel, G. The activity of siRNA in mammalian cells is related to structural target accessibility: a comparison with antisense oligonucleotides. *Nucleic Acids Res.* **31**, 4417–4424 (2003).
46. Lima, W. F., Monia, B. P., Ecker, D. J. & Freier S. M. Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry* **31**, 12055–12061 (1992).
47. Yoshinari, K., Miyagishi, M. & Taira, K. Effects on RNAi of the tight structure, sequence and position of the targeted region. *Nucleic Acids Res.* **32**, 691–699 (2004).
48. Frantz, S. Studies reveal potential pitfalls of RNAi. *Nature Rev. Drug Discov.* **2**, 763–764 (2003).
49. Monia, B. P. *et al.* Selective inhibition of mutant Ha-ras mRNA expression by antisense oligonucleotides. *J. Biol. Chem.* **267**, 19954–19962 (1992).
50. Wahlestedt, C. *et al.* Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl Acad. Sci. USA* **97**, 5633–5638 (2000).
51. McCaffrey, A. P., Meuse, L., Karimi, M., Contag, C. H. & Kay, M. A. A potent and specific morpholino antisense inhibitor of hepatitis C translation in mice. *Hepatology* **38**, 503–508 (2003).
52. Pancoska, P., Moravek, Z. & Moll, U. M. Efficient RNA interference depends on global context of the target sequence: quantitative analysis of silencing efficiency using Eulerian graph representation of siRNA. *Nucleic Acids Res.* **32**, 1469–1479 (2004).
53. Freier, S. M. in *Antisense Research and Applications* (Eds Crooke, S. T. & Lebleu, B.) Ch. 5 (CRC Press, Boca Raton, Florida, 1993).
54. Donze, O. & Picard, D. RNA interference in mammalian cells using siRNAs synthesized with T7 RNA polymerase. *Nucleic Acids Res.* **30**, e46 (2002).
55. Miyagishi, M. & Taira, K. U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nature Biotechnol.* **20**, 497–500 (2002).
56. Lee, N. S. *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nature Biotechnol.* **20**, 500–505 (2002).
57. Nykanen, A., Haley, B. & Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309–321 (2001).
58. Djikeng, A., Shi, H., Tschudi, C., Shen, S. & Ullu, E. An siRNA ribonucleoprotein is found associated with polyribosomes in *Trypanosoma brucei*. *RNA* **9**, 802–808 (2003).
59. Ishizuka, A., Siomi, M. C. & Siomi, H. A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev.* **16**, 2497–2508 (2002).

Acknowledgements

I thank D. Huesken, J. Hunziker and D. Garza for useful discussions, and W. Filipowicz and D. Cohen for reading the manuscript.

Competing interests statement

The author declares that he has no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:

Entrez: http://www.ncbi.nih.gov/Entrez/

AKT1 | *CD54* (*ICAM1*) | *CXCR4* | *Gag* | *geminin* (*GMNN*) | *IGF1R* | *Lmna* | *MEN1* | *MJD* | *p21* (*CDKN1A*) | *p53* (*TP53*) | *PLK1* | *PTEN* | *RB1*

Access to this links box is available online.