

A rapid and sensitive assay for quantification of siRNA efficiency and specificity

Nicola Smart¹, Peter James Scambler¹ and Paul Richard Riley^{1*}

¹Molecular Medicine Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, United Kingdom.

*Corresponding Author: Dr. Paul R. Riley, Molecular Medicine Unit, 30 Guilford Street, London WC1N 1EH. Phone: 020 7905 2345; Fax: 020 7404 6191; Email: P.Riley@ich.ucl.ac.uk

Submitted: October 14, 2004; Revised: December 21, 2004; Accepted: December 23, 2004.

Indexing terms: Thymosin beta(4); Thymosin beta(10).

ABSTRACT

RNA Interference has rapidly emerged as an efficient procedure for knocking down gene expression in model systems. However, cross-reactivity, whereby multiple genes may be simultaneously targeted by a single short interfering RNA (siRNA), can potentially jeopardize correct interpretation of gene function. As such, it is essential to test the specificity of a siRNA prior to a full phenotypic analysis. To this end, we have adapted a reporter-based assay harnessing the sensitivity of luciferase activity to provide a quantitative readout of relative RNAi efficacy and specificity. We have tested different siRNAs directed against *Thymosin β4* (*Tβ4*); determined their effectiveness at silencing *Tβ4* and have both excluded off-target silencing of the *Tβ4* homologue *Thymosin β10* (*Tβ10*) and demonstrated partial knockdown of *Tβ10* despite significant (12/23; 52%) sequence mismatch. This assay system is applicable to any RNAi study where there is a risk of targeting homologous genes and to the monitoring of off-target effects at the genome level following microarray expression profiling.

INTRODUCTION

The ability to selectively abolish expression of a single gene and to examine the resulting cellular phenotype has enabled substantial progress in the understanding of gene function. Despite the reported problems of cross-reactivity (1) and induction of an interferon response (2), RNA interference (RNAi) has rapidly become the preferred technique for abolishing gene expression.

RNAi, which occurs endogenously in eukaryotes, is initiated by the cleavage of long dsRNA into 21-23 nucleotide siRNA, by the RNase III endonuclease, Dicer. The siRNAs are incorporated into a multisubunit RNA-induced silencing complex (RISC), which targets complementary mRNA for degradation.

The extent of target specificity offered by RNAi in mammalian cells is still highly controversial and is

almost certainly sequence-dependent. Many groups report a high degree of specificity, such as the ability of siRNAs to discriminate between wild type and mutant p53, differing by a single nucleotide, in H1299 and U20S cells (3). Genome-wide microarray studies have attempted to resolve the matter, but have proven equally inconclusive (1, 4-6). Specific silencing was observed when an exogenously expressed protein, EGFP, was targeted in HEK 293 cells (5). However, a conflicting study in HeLa cells demonstrated that RNAi targeting two different genes, IGF1R and MAPK14, resulted in silencing of non-targeted genes containing as few as eleven nucleotides of identity (1). Disparity between these two data sets could be explained by the fact that EGFP is not an endogenous mammalian gene and, therefore, siRNAs directed towards IGF1R and MAPK14 would be expected to share sequence homology with a greater number of mammalian genes than a siRNA to target EGFP and carry a greater risk of off-target

silencing. Either way it appears that there are intrinsic limits on the specificity of RNAi. In a recent study examining 359 published siRNA sequences, it was found that approximately 75% have a risk of non-specific targeting (7). Furthermore, siRNAs can illicit non-specific effects at the level of protein synthesis without a corresponding effect on mRNA levels (8) in a manner analogous to that observed with micro RNAs (miRNAs) (9-10).

Clearly, so-called off-target gene silencing could represent a significant problem, resulting in RNAi adversely targeting multiple genes simultaneously. It is, therefore, essential to determine which genes, if any, share significant sequence identity with the proposed siRNA and to directly assess whether their expression is altered by the siRNA.

Our RNAi studies have been based upon the targeting of the actin binding protein Thymosin beta 4 (*Tβ4*). We have encountered significant difficulties, such as probe cross-reactivity and non-specific product detection, using conventional methods, in assessing the specificity of *Tβ4* knockdown as against a non-specific effect on its closest relative Thymosin beta 10 (*Tβ10*). This is a potential problem associated with detecting expression levels following RNAi of many similarly sized, highly homologous genes and/or members of the same gene family.

In this study we present a reporter-based assay system to test both the efficacy and specificity of *Tβ4* knockdown, which can be applied in general to other siRNAs. We have placed target sequences for either *Tβ4* or *Tβ10* downstream of a luciferase gene and following co-transfection of the relevant siRNA assessed specific knockdown via a quantitative reduction in luciferase activity.

MATERIALS AND METHODS

Construction of *Tβ4* shRNAs

Tβ4 shRNA constructs were prepared by modifying a RasGAP shRNA transgene (11), kindly provided by G.Gish, which uses the human H1 RNA polymerase III to drive transcription of short hairpin RNA (shRNA) targeting the p120-Ras GTPase-activating protein (*RasGAP*). The RasGAP shRNA sequence was removed

from the transgene by digest with *Asp718* and *XbaI*, to leave a pcDNA3.1 (+) backbone (Invitrogen) containing the pol III promoter. The target *Tβ4* shRNA sequences, underlined in Figure 1, were prepared by oligonucleotide synthesis: 5'-GTACCCTGAGATCGAGAAATTCGATAAGTTCAAGA GACTTATCGAATTTCTCGATCTCAGTTTTTGGAAAT-3' (*Tβ4* shRNA1, sense strand); 5'-CTAGATTTCCAAAACTGAGATCGAGAAATTCGAT AAGTCTCTTGAACCTTATCGAATTTCTCGATCTCAGG-3' (*Tβ4* shRNA1, antisense strand); 5'-GTACCAGAAGCAAGCTGGCGAATCGTAATTCAAG AGATTACGATTCGCCAGCTTGCTTCTTTTTGGAAA T-3' (*Tβ4*, shRNA2, sense strand); 5'-CTAGATTTCCAAAAAGAAGCAAGCTGGCG AATCGTAATCTCTTGAATTACGATTCGCCAG CTTGCTTCTG-3' (*Tβ4* shRNA2, antisense strand). Oligonucleotides were annealed by heating to 90°C in the presence of 50mM NaCl and gradual cooling to room temperature. Double stranded oligonucleotides were introduced into the *Asp718* and *XbaI* sites of pcDNA3-H1 pol III.

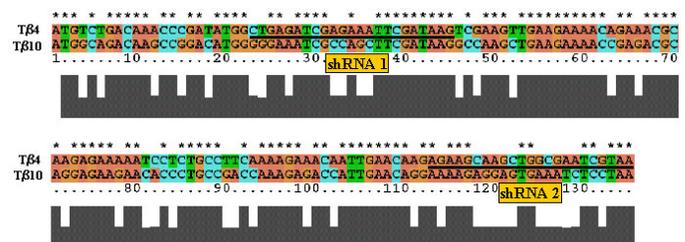


Fig. 1: *Tβ4* and *Tβ10* share a high degree of sequence homology. An alignment of *Tβ4* and *Tβ10* coding sequences, using ClustalW (MRC RFCGR, <http://www.rfcgr.mrc.ac.uk/Registered/Option/clustal.html>). Sequences selected for *Tβ4* shRNA constructs, on the basis of lower homology with *Tβ10*, are underlined.

Generation of β -thymosin reporter constructs

Oligonucleotides were synthesised to include the 23nt target sequence along with 38nt of flanking sequence: 5'-CAAGGATGCTGACAAACCCGATATGGCTGAGATC GAGAAATTCGATAAGTTCGAAGTTGAAGAAAACAGAA C-3' (*Tβ4* reporter 1, sense strand); 5'-CTTGGTTCGTGTTTTCTTCAACTTCGACTTATCGAATTC TC GATCTCAGCCATATCGGGTTTGTGACACATC-3' (*Tβ4* reporter 1, antisense strand); 5'-CAAGGATGGCAGACAAGCCGGACATGGGGGAAATCGCA GCIT CGATAAGGCCAAGTTGAAGAAAACCGAGC-3' (*Tβ10* reporter 1, sense strand); 5'-CTTGGCTCGGTTTTCTTCAACTTGGCCTTATCGAAGCT

GGCGATTCCCC CATGTCCGGCTTGTCTGCCATC-3' (T β 10 reporter 1, antisense strand); 5'-CAAGGTCAAAGAAACAATTGAACAAGAGAAGCAA GCTGGCGAATCGTAATGAGGCGAGCGCCGCAATAT C-3' (T β 4 reporter 2, sense strand); 5'-CTTGATATTGGCGGCGCTCGCCTCATTACGATTCCGC AGCTTGCTTCTTGTCAATTGTTCTTTTGAC-3' (T β 4 reporter 2, antisense strand); 5'-CAAGGACCAAAGAGACCATTGAACAGGAAAAGAGG AGTGAAATCTCCTAAAAGCCTAGGAAGATTTCCCCAC -3' (T β 10, reporter 2, sense strand); 5'-CTTGGTGGGAAATCTCCTAGGCTTTTAGGAGATTC ACTCCTCTTTTCTGTTCAATGGTCTCTTTGGTC-3' (T β 10 reporter 2, antisense strand). Oligonucleotides were annealed as described above and ligated into the *StyI* site of pGL2-Basic (Promega). To enable a high level of basal expression, the SV40 promoter of pGL2-Control (Promega) was subcloned as a *Hind III/Kpn I* fragment into the pGL2-Basic polylinker, upstream of the target sequences.

Culture and transfection of NIH-3T3 cells

All cell culture reagents were from Invitrogen. Cells were maintained in DMEM, supplemented with 10% (w/v) bovine calf serum, 100units/ml penicillin and 100 μ g/ml streptomycin, at 37°C in a humidified 5% CO₂/air atmosphere. Cells were transfected with the shRNA or co-transfected with shRNA and reporter constructs, using Effectene (Qiagen) transfection reagent. 2.1 x 10⁵ cells were cultured overnight to 50-70% confluence and a total of 500ng DNA added per well. A pCMV- β -gal plasmid was also included as a marker of transfection efficiency. Duplicate transfections were performed in each experiment and data pooled from four independent experiments.

Luciferase reporter assays

Cells were harvested 48 hours post-transfection and assayed using Promega's Luciferase Assay System. Cells were washed once in PBS, scraped into 400 μ l of reporter lysis buffer and lysis facilitated by two rounds of freezing (-80°C) and thawing (37°C). 20 μ l lysate were combined with 100 μ l assay reagent and luciferase

activity measured using a TD-20/20 Luminometer (Turner Designs). To normalize for transfection efficiency, the activity of co-transfected β -galactosidase was also assayed. 100 μ l of lysate were combined with 50 μ l of lysis buffer and 150 μ l of 2x β -galactosidase assay buffer and incubated at 37°C for 20-90 minutes until the yellow o-nitrophenyl was produced. The reaction was terminated by the addition of 1M Na₂CO₃ and the absorbance at 420nm recorded.

RESULTS AND DISCUSSION

We sought to target *Thymosin β 4*, which encodes an actin-binding protein implicated in cell movement, using RNAi. *T β 4*, the predominant mammalian isoform of the β -thymosin family, shares a high degree of homology with other isoforms, such as *T β 10* and *T β 9* (12). *T β 10* is the β -thymosin most closely related to *T β 4*, displaying 70.4% identity within the coding region. The coding region for both *T β 4* and *T β 10* is only 135bp and as such there is limited potential to design a *T β 4*-specific siRNA that will not cross-react with *T β 10*. Intronic sequences are ordinarily avoided since these immature mRNAs will be kept in the nucleus while RNAi mediated mRNA degradation is believed to occur predominantly in the cytoplasm (13). 3' and 5' untranslated regions tend to be avoided since they may be bound by regulatory proteins and thus be inaccessible to siRNAs or RISC. Further limitations are imposed by the need to adopt established criteria for optimal siRNA design (14).

Sense and antisense *T β 4* sequences of 23 base pairs in length, separated by a nine bp spacer, were introduced downstream of the H1 RNA pol III promoter (11, 15), followed by a stretch of five thymidines which act to terminate transcription. Two sequence stretches were selected, as regions displaying minimum sequence identity with *T β 10*; T β 4shRNA1 incorporated six mismatches and T β 4shRNA2 twelve mismatches versus *T β 10*, as shown in Figure 1. The mismatches lie predominantly in the center of the shRNAs, with additional mismatches at the 5' end of T β 4shRNA1 and at both extremes of T β 4shRNA2. The position of a mismatch has been shown to be a critical factor in RNAi selectivity, with mismatches at the 5' end more readily

tolerated than those at the 3' end (16). Given the number and positions of the mismatches, neither T β 4shRNA would be predicted to cross-react with T β 10. Constructs were evaluated by transient transfection of NIH-3T3 cells. We encountered considerable difficulty when attempting to demonstrate the specificity of T β 4 RNAi. While northern blotting appeared to show down-regulation of T β 4 (to 21% of the level in non-transfected cells), re-probing of the blot for T β 10 levels revealed bands of equivalent size and signal intensities as seen for T β 4. Due to the significant level of identity shared by these two genes, the T β 10 probe was able to cross-react with the more abundant T β 4 transcript, which is comparable in size. Quantitative Real Time RT-PCR proved equally uninformative; although specific products could be amplified for T β 4 and T β 10, the degree of T β 4 knockdown revealed by this technique was modest, likely due to the sensitivity in detecting expression in cells which failed to transfect with the shRNA constructs (estimated to be as high as 40%).

The need for a more sensitive means of determining RNAi specificity led us to develop a reporter-based assay for RNAi, as outlined in Figure 2. Our assay is based upon the introduction of target sequences downstream of the luciferase reporter gene. The ability of a complementary shRNA to knockdown its target is then quantitatively determined by a decrease in luciferase activity. Synthetic oligonucleotides containing the 23nt target region along with 40nt of flanking sequence were designed to produce the appropriate *StyI* overhangs when annealed (Fig. 2). Oligonucleotides were ligated into the *StyI* site located between the splice donor site and poly(A) signal of the luciferase gene in pGL2-Basic (Promega), resulting in the transcription of a luciferase mRNA containing the target sequence. The non-palindromic nature of the *StyI* site in pGL2-Basic (CCAAGG) ensured that the oligonucleotides were ligated into the vector in the desired 5'-3' orientation. The SV40 promoter of pGL2-Control (Promega) was introduced as a *KpnI*/*HindIII* fragment into the multiple cloning region, to enable strong basal luciferase expression.

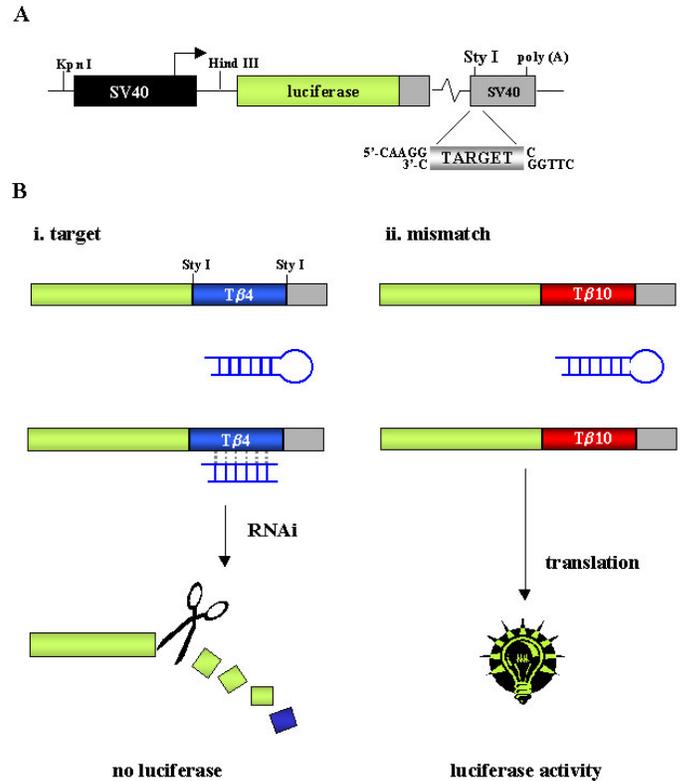


Fig. 2: A reporter assay to determine the specificity of RNAi. A. Oligonucleotides were synthesized to contain the 23nt shRNA target region, 20nt of flanking sequence on each side and appropriate nucleotides to generate *StyI* overhangs. Complementary oligonucleotides were annealed and ligated into the *StyI* site of pGL2-Basic. The SV40 promoter of pGL2-Control was subcloned into the *KpnI*/*HindIII* sites for maximal expression. **B.** The use of the reporter assay to determine whether a gene can be specifically targeted for RNAi (i) and to confirm that a closely related gene remains unaffected (ii). The respective T β 4 shRNA target sequences and analogous sequences of T β 10 were introduced into reporter constructs as described above. Upon transfection into mammalian cells luciferase mRNA is transcribed containing the target sequence within the 3' untranslated region. Translation is unaffected by the presence of the target sequence and functional luciferase is synthesized. T β 4 shRNAs, when co-transfected and transcribed by H1 polymerase III, direct the RISC towards the complementary T β 4 target sequence within the luciferase mRNA. Degradation of the mRNA ensues, translation does not occur and no luciferase activity is detected. A mismatched target sequence (T β 10) in principal is not recognized by the shRNA, the mRNA is translated and luciferase activity detected.

Four reporter constructs were created, two of which contained the T β 4 shRNA target regions and the other two contained the analogous regions of T β 10 (see alignment, Fig. 1). Co-transfection of T β 4 reporter 1 with the appropriate shRNA (T β 4shRNA1) resulted in a significant knockdown of luciferase activity (by 87%, n=8, p<0.0001) compared with the pcDNA3 backbone, confirming the efficacy of the shRNA (Fig. 3A). Neither of the mismatched shRNAs (T β 4shRNA2 nor a control RasGAP shRNA (11)) affected luciferase activity.

Conversely, activity of T β 4 reporter 2 was selectively knocked down by T β 4shRNA2 (by 89%, n=8, p<0.0001), but not by T β 4shRNA1 (Fig. 3B). Activity of T β 10 reporter 1 was unaffected by the T β 4shRNA1 as compared to the RasGAPshRNA control (Fig. 3A) and neither, T β 4shRNA1 nor T β 4shRNA2 adversely affected T β 10 reporter 2 (Fig. 3B). However, the T β 4shRNA2 construct, which contained 12 sequence mismatches to the analogous T β 10 region, reduced T β 10 reporter 1 activity (by 42%, n=8, p<0.001; Fig. 3A) compared to the pcDNA3 backbone. Given the degree of sequence mismatch this was an unexpected, yet significant observation since it represents partial silencing of an undesired “target.” This in turn could potentially result in a hypomorphic cellular phenotype assuming a dose-dependent requirement for T β 10 function.

The reporter assay, therefore, has enabled us to compare the potency of shRNAs in inducing RNAi against *T β 4* and, importantly, to test the selectivity of knockdown of each shRNA against *T β 10*. This method can be applied to all RNAi-based studies in which there is risk of targeting homologous genes, particularly when selective knockdown has been difficult to prove by conventional semi-quantitative assays of gene expression.

The present study exposes a further issue pertaining to RNAi, which is that siRNAs do not have to target coding region; although this is usually the preferred approach (13, 17). Whilst our shRNAs are directed against the coding sequence of the gene of interest, the sequence has been inserted into the 3'UTR of the luciferase gene, where the shRNAs induce degradation of the luciferase mRNA. In cases where sequence homology is high throughout the coding region and off-target silencing is shown to occur, it may become necessary to target the more divergent untranslated sequence. It should be possible to determine, using reporter constructs based on those described herein, whether binding by regulatory proteins renders the sequence inaccessible to the RNAi machinery upon transfection into the cell type for phenotype analysis.

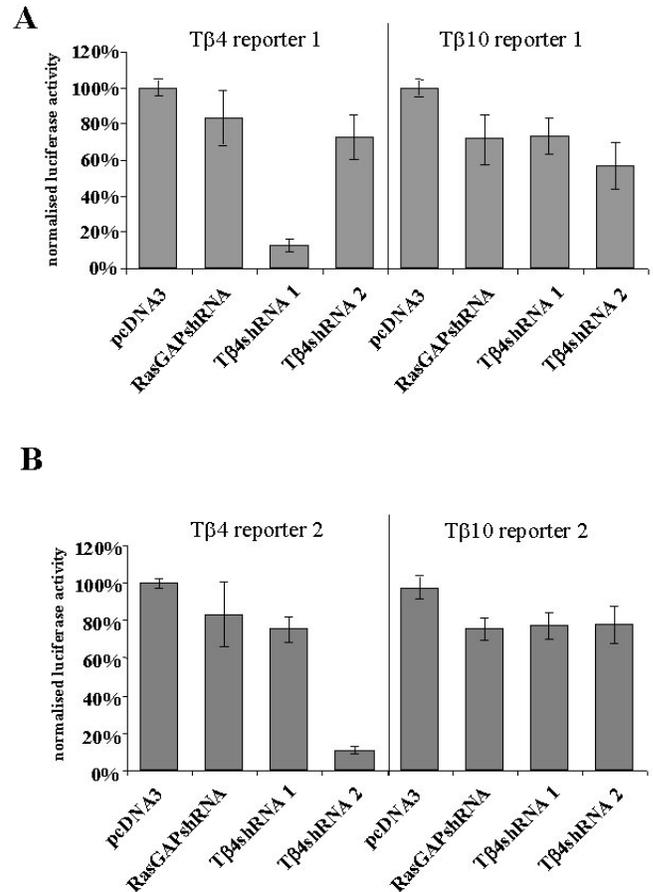


Fig. 3: Use of the luciferase reporter assay to determine the specificity of T β 4 shRNA. NIH3T3 cells were transfected with a reporter construct, containing either of the *T β 4* target regions or the analogous regions of *T β 10* (n=8; duplicates from four independent experiments). The ability of the T β 4 shRNAs to induce RNAi was assessed by co-transfection; T β 4shRNA1, the shRNA complementary to T β 4 reporter 1, induced an 87% down-regulation of luciferase activity (p<0.0001), compared with empty pcDNA3 control vector (A). Non-complementary shRNAs (T β 4shRNA2 or a control RasGAPshRNA (11) did not significantly reduce luciferase activity. Similarly, T β 4shRNA2 selectively induced an 89% reduction of the luciferase activity of T β 4 reporter 2 (p<0.0001, B) The specificity of T β 4 RNAi derived from T β 4shRNA1 was confirmed by transfection of either of the T β 10 reporter constructs, the expression of which was not significantly decreased.

The use of a reporter gene assay offers a rapid and reliable means of quantitatively assessing the potency and specificity of RNAi. With the approach described here, the generation of constructs is straightforward, relying on the use of synthetic oligonucleotides and a single restriction site that ensures directional cloning. Assaying reporter activity requires little effort and unequivocal data can be rapidly obtained, enabling immediate progression to the next experimental stage. Not all siRNAs are equally effective in inducing RNAi and much optimization is often required. Using a reporter approach, multiple siRNAs, directed towards

the same gene, may be simultaneously assayed and quantitatively compared.

The caveats in specificity of gene silencing by RNAi means there is an absolute requirement to test the selectivity of a siRNA before embarking on phenotype analysis. Previous use of reporter assays, with respect to RNAi, has enabled the testing of general knockdown (18-20), but the method has not been applied as a test for specificity. The comparison in this study of two closely related β -thymosin genes is testament to the sensitivity of this approach for determining RNAi specificity. The observation that one of our T β 4 shRNA constructs induced 42% knock-down of a T β 10 reporter, despite significant mismatches in sequence, further underscores the need for an assessment of the specificity of multiple shRNAs. Despite the fact that the gene most closely related to T β 4 was unaffected by the T β 4shRNA1 construct, we cannot fully exclude the possibility that there are other genes in the genome which can be silenced by this construct. That said, use of this assay in conjunction with optimal siRNA design, the testing of several different siRNAs for equivalent phenotypic outcome and identification of RNAi triggers at lowest possible concentration (21) should significantly improve confidence in the specificity of gene knockdown.

ACKNOWLEDGMENTS

This work was supported by Project Grants PG/2000117 and PG/04/012/16626 from the British Heart Foundation.

REFERENCES

- Jackson A, Bartz S, Scheller J, Kobayashi S, Burchard J, Mao M, Li B, Cavet G, Linsley P. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology* 2003; 21:635-637.
- Bridge A, Pebernard S, Ducraux A, Nicoulaz A-L, Iggo R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nature Genetics* 2003; 34:263-264.
- Martinez LA, Naguibneva I, Lehrmann H, Vervisch A, Tchenio T, Lozano G, Harel-Bellan A. Synthetic small inhibiting RNAs: Efficient tools to inactivate oncogenic mutations and restore p53 pathways. *PNAS* 2002; 99:14849-14854.
- Semizarov D, Frost L, Sarthy A, Kroeger P, Halbert DN, Fesik SW. Specificity of short interfering RNA determined through gene expression signatures. *PNAS* 2003; 100:6347-6352.
- Chi JT, Chang HY, Wang NN, Chang DS, Dunphy N, Brown PO. Genomewide view of gene silencing by small interfering RNAs. *PNAS* 2003; 100:6343-6346.
- Persengiev SP, Zhu XIAO, Green MR. Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). *RNA* 2004; 10:12-18.
- Snove J, Holen T. Many commonly used siRNAs risk off-target activity. *Biochemical and Biophysical Research Communications* 2004; 319:256-263.
- Scacheri PC, Rozenblatt-Rosen O, Caplen NJ, Wolfsberg TG, Umayam L, Lee JC, Hughes CM, Shanmugam KS, Bhattacharjee A, Meyerson M, Collins FS. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *PNAS* 2004; 101:1892-1897.
- Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of Mammalian MicroRNA Targets. *Cell* 2003; 115:787-798.
- Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome* 2003; Biol 5:R1.
- Kunath T, Gish G, Lickert H, Jones N, Pawson T, Rossant J. Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. *Nature Biotechnology* 2003; 21:559-561.
- Huff T, Muller C, Otto A, Netzker R, Hannappel E. β -thymosins, small acidic peptides with multiple functions. *Int J Biochem & Cell Biol* 2001; 33:205-220.
- Hannon G. RNA Interference. *Nature* 2002; 418:244-251.
- Cui W, Ning J, Naik UP, Duncan MK. OptiRNAi, an RNAi design tool. *Computer Methods and Programs in Biomedicine* 2004; 75:67-73.
- Myslinski E, Ame JC, Krol A, Carbon P. An unusually compact external promoter for RNA polymerase III transcription of the human H1RNA gene. *Nucl Acids Res* 2001; 29:2502-2509.
- Amarzguoui M, Holen T, Babaie E, Prydz H. Tolerance for mutations and chemical modifications in a siRNA. *Nucl Acids Res* 2003; 31:589-595.
- Elbashir S, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001; 411:494-498.

18. Kumar R, Conklin DS, Mittal V. High-Throughput Selection of Effective RNAi Probes for Gene Silencing. *Genome Res* 2003; 13:2333-2340.
19. Zeng Y, Yi R, Cullen BR. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *PNAS* 2003; 100:9779-9784.
20. Holen T, Amarzguioui M, Wiiger MT, Babaie E, Prydz H. Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucl Acids Res* 2002; 30:1757-1766.
21. Hannon GJ, Rossi JJ. Unlocking the potential of the human genome with RNA interference. *Nature* 2004; 431:371-378.