Published online February 3, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 2 ***691–699***
DOI: 10.1093/nar/gkh221

Effects on RNAi of the tight structure, sequence and position of the targeted region

Koichi Yoshinari[1], Makoto Miyagishi[1,2] and Kazunari Taira[1,2,*]

[1]Gene Function Research Center, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba Science City 305-8562, Japan and [2]Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Hongo, Tokyo 113-8656, Japan

Received June 15, 2003; Revised September 9, 2003; Accepted December 17, 2003

ABSTRACT

RNA interference (RNAi) is a gene-silencing phenomenon that involves the double-stranded RNA-mediated cleavage of mRNA, and small interfering RNAs (siRNAs) can cause RNAi in mammalian cells. There have been many attempts to clarify the mechanism of RNAi, but information about the relationship between the sequence and structure, in particular, a tight structure, of the target RNA and the activities of siRNAs are limited. In the present study, we examined this relationship by introducing the TAR element, which adopts a very stable secondary structure, at different positions within target RNAs. Our results suggested that the activities of siRNAs were affected by the tight stem–loop structure of TAR. In contrast, the position of the target within the mRNA, the binding of the Tat protein to the TAR, and the location of the target within a translated or a noncoding region had only marginal effects on RNAi. When the target sequence was placed in two different orientations, only one orientation had a significant effect on the activities of siRNA, demonstrating that the presence of certain nucleotides at some specific positions was favorable for RNAi. Systematic analysis of 47 different sites within 47 plasmids under identical conditions indicated that it is the target sequence itself, rather than its location, that is the major determinant of siRNA activity.

INTRODUCTION

RNAi (RNA interference) is a gene-silencing phenomenon that involves the double-stranded RNA (dsRNA)-mediated cleavage of mRNA. It has been demonstrated in plants, nematodes, *Drosophila*, protozoa and mammalian cells (1–8). In RNAi, dsRNA is cleaved into small RNAs of ~21–25 nt in length, referred to as small interfering RNAs (siRNAs). Dicer, a member of the ribonuclease III family (9–15), catalyzes these endonucleolytic cleavages. The siRNAs form a multicomponent nuclease complex known as an RNA-induced silencing complex (RISC) (16–19) and the siRNAs function as guide RNAs, directing the complex to the target mRNA (17,19). The targeted mRNA is recognized by protein factors within the RISC and cleaved by the action of a nuclease within the RISC (16,18).

RNAi has been exploited as a powerful tool in reverse-genetic studies of *Caenorhabditis elegans* (20,21), and specific and efficient RNAi has been achieved in mammalian cells with duplexes of 21 nt RNAs that form a 19 bp region with a 2 nt 3′ overhang (4,6,22–24). Such a duplex is small enough to bypass the interferon responses of differentiated cultured cells (23). Thus, RNAi has the potential to become a valuable tool for analysis of the biological functions of genes in mammalian cells.

It is now possible to generate siRNAs in cells from various expression vectors and, therefore, both synthetic and vector-derived siRNAs can be used in the functional analysis of genes of interest in mammalian cells (12,25–33). Several studies exploiting RNAi have been performed (8). Although the dependence of the activities of siRNAs on their target sequences was reported recently (26,27,34–37), the selection of the best target site is often difficult (12). The published data suggest that the GC content of the target sequence, the position and accessibility of the target site, and the strength of terminal base-pairings of siRNAs are important.

In this study, to analyze in greater detail the parameters that govern the activities of siRNAs, we constructed a series of structured target RNAs and systematically analyzed the extent of RNAi in the presence and in the absence of a protein that can interact with the target RNAs.

MATERIALS AND METHODS

Construction of plasmids

In most experiments (Fig. 1), we used the vector pTAR-GL2, which contains a TAR motif (Fig. 1B) upstream of the initiation codon of the firefly luciferase gene. To compare the effects of position of the TAR motif and other target sequences on RNAi, we constructed derivatives of pGL3 (pGL3-control vector; Promega, Madison, WI; accession no. U47296) that

*To whom correspondence should be addressed at Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Hongo, Tokyo 113-8656, Japan. Tel: +81 3 5841 8828 or +81 29 861 3015; Fax: +81 3 5841 8828 or +81 29 861 3019; Email: taira@chembio.t.u-tokyo.ac.jp

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

 © Oxford University Press 2004; all rights reserved

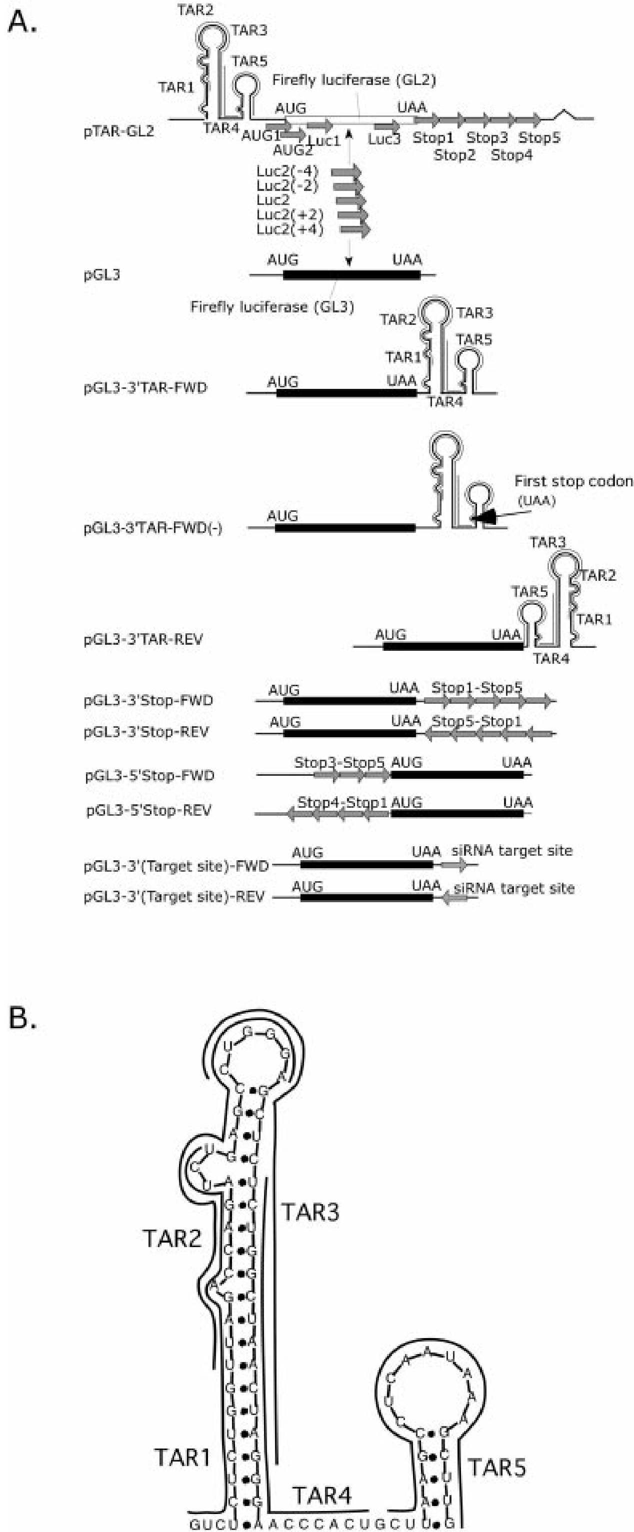

Figure 1. The reporter mRNAs for firefly luciferases that were used in the present study. (**A**) The white box represents the firefly luciferase gene GL2 and the black boxes represent the firefly luciferase gene GL3. TAR motifs are shown with their putative secondary structures at their respective positions in each gene. The fat gray arrows indicate the target sites of siRNAs. The orientation of the arrows indicates the direction of each sequence within the mRNA. (**B**) The target sites corresponding to the various TAR motifs are shown by thin lines.

had a TAR element downstream of the stop codon of the firefly luciferase gene or that had the sequence that is normally located around the stop codon of the firefly luciferase gene in pTAR-GL2 upstream of the initiation codon of firefly luciferase gene or downstream of the stop codon of this gene (Fig. 1A).

The vector with a TAR motif downstream of the stop codon of the firefly luciferase in pGL3 was constructed as follows. TAR and the sequential LTR sequence of pTAR-GL2 (TAR motif) were amplified by PCR with primers that included an XbaI site, namely, 5′-GCG CTC TAG AGG GTC TCT CTG GTT AGA-3′ and 5′-GGG CTC TAG ATG CCA AGC TTT ATT GAG G-3′. The amplified fragment was ligated into the XbaI site of pGL3 and two kinds of vector were obtained, one had the TAR motif inserted in the forward orientation (pGL3-3′TAR-FWD) and the other had the TAR motif in the reverse orientation (pGL3-3′TAR-REV).

We also constructed a vector, pGL3-3′TAR-FWD(–), by deleting two stop codons from the pGL3-3′TAR-FWD. First, the original stop codon, TAA, of firefly luciferase gene was changed to AAA, and then the TAG triplet, located in the first XbaI site, was changed to AAG. Sequences between the SgrAI site and the FseI site of pGL3-3′TAR-FWD were amplified by PCR with two sets of primers, GTG TCG CAG GTC TTC CCG plus CCA GAG AGA CCC TCT **T**GA ATT **T**CA CGG CGA TCT TTC C, and GGG AAG ATC GCC GTG **A**AA TTC **A**AG AGG GTC TCT CTG G plus TCT TAT CAT GTC TGC TCG AAG (bold letters show mutations that were introduced to eliminate stop codons). The amplified fragments were combined by further PCR and ligated into the SgrAI site and the FseI site of pGL3. In this construct, the first stop codon, TAA, is located within target site TAR5 (this target site is outside the stem–loop structure of TAR).

The vector with the stop-codon region of the firefly luciferase gene of pTAR-GL2 upstream of the initiation codon of the firefly luciferase gene in pGL3 was constructed as follows. The sequence that included the stop codon of pTAR-GL2 was amplified with two sets of primers, which contained a HindIII and an NcoI site, respectively: FWD (5′-CTG AAG CTT AAT ACT CTA GAG GAT CTT TGT-3′ plus 5′-CGG TCC ATG GTA GGT AGT TTG TCC AAT TAT-3′) and REV (5′-GAC AAG CTT ACA CCA CAG AAG TAA GGT TCC-3′ plus 5′-CGG CCC ATG GTT GTA AAA TGT AAC TGT ATT-3′). The amplified fragments obtained with primer set FWD and primer set REV were ligated into the HindIII site and the NcoI site of pGL3 and newly constructed vectors were designated pGL3-5′Stop-FWD and pGL3-5′Stop-REV, respectively.

The stop-codon region of pTAR-GL2 was also introduced after the stop codon of the firefly luciferase gene of pGL3 as follows. The stop-codon region of pTAR-GL2 was amplified with primers that contained an SpeI site, namely, 5′-CTG ACT AGT TGT AAA ATG TAA CTG TAT TCA-3′ and 5′-CGG ACT AGT TAG GTA GTT TGT CCA ATT AT-3′. The amplified fragment was ligated into the XbaI site of pGL3, and two kinds of vector were obtained, one having the insert in the forward direction (pGL3-3′Stop-FWD) and one having the insert in the reverse direction (pGL3-3′Stop-REV).

We also constructed derivatives of pGL3 with 23mer target sequences after the stop codon of the firefly luciferase gene. The DNAs with each siRNA target sequence (sense

and antisense sequence: CTAGT-N_{23}-A) were synthesized commercially (Proligo, Kyoto, Japan). A mixture of sense and antisense DNAs was phosphorylated with T4 polynucleotide kinase. After incubation at 37°C for 1 h, the enzyme was inactivated by heating at 95°C for 1 min and DNAs were allowed to anneal at room temperature. The various fragments were ligated separately into the XhoI site of pGL3.

Preparation of siRNAs

Each strand of siRNAs was synthesized chemically, by Japan Bio Service (Saitama, Japan) as a 19mer RNA and as 3′-dimeric DNA chimeras. It has been reported that the 5′-OH group must be phosphorylated before formation of RISC can occur (19) and that the activity of 5′-OH siRNA is lower than that of 5′-phosphorylated siRNA (38). Therefore, we used siRNAs that had been 5′-phosphorylated chemically. The sequences of the siRNAs used in this study (Figs 2–5) are shown in Table 1. The annealing of sense and antisense strand RNA–DNA chimeras was performed as described previously (26). We examined the annealing efficiency of 5′-phosphorylated sense and antisense strands under the previously described conditions using 5′ ^{32}P-labeled sense and antisense RNA–DNA chimeras targeted to the TAR motif (for this experiment, the synthetic RNAs had 5′-OH groups). As shown in Figure 2D, the annealing efficiencies for these siRNA were at least 77%. Even TAR1, of which each strand alone was able to form a stem–loop structure, yielded nearly 96% siRNA.

Introduction of siRNAs by transfection and the dual luciferase assay

The RNAi assay was performed under the conditions described previously (26). We cultured HeLa S3 cells (3×10^4 cells/well) in 48-well plates for 24 h before transfection with the firefly luciferase expression vector (pTAR-GL2, pGL3 or derivatives of pGL3), the Tat expression vector (pCD-SRα/Tat), the *Renilla* luciferase expression vector (pRL-RSV), and annealed siRNA, using Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA). The final amount of each vector was 25 ng in 250 µl and the final concentration of siRNA was 1 nM. After further incubation for 24 h, the activities of the firefly luciferase and the *Renilla* luciferase were measured with a dual luciferase assay kit (Promega) and a Lumat detection system (LB9501; Berthold, Bad Wildbad, Germany).

RESULTS AND DISCUSSION

Design of plasmids

To examine the effects on RNAi of the tight structure, position and sequence of a target RNA and of binding of a peptide to such a target, we constructed 47 plasmids, as shown schematically in Figure 1A. To investigate the effects of both tight structure and the binding of a peptide to the target, we chose the TAR motif. To examine positional effects, we placed the TAR motif on either the 5′ or the 3′ side of a firefly luciferase gene, generating pTAR-GL2 and pGL3-3′TAR-FWD, respectively. To examine the effects of the tight structure of the TAR motif, we flipped the TAR sequence within pGL3-3′TAR-FWD to yield pGL3-3′TAR-REV, in which the nonfunctional flipped target motif retained the original tight structure of the TAR motif. If the tight structure itself, but not the binding of Tat protein to TAR, were to govern the activity of siRNA, the nonfunctional 'flipped' TAR derived from pGL3-3′TAR-REV should give results similar to those obtained with pGL3-3′TAR-FWD.

To distinguish the effects of sequence from the effects of structure, we isolated the target sequence from the TAR motif (TAR1-TAR5 in Fig. 1B), as well as all the other target sequences (Stop1–Stop5, ATG1–ATG2, Luc1–Luc3) used in this study, and placed each individual target sequence in the same 3′-untranslated region, in the forward and, separately, in the reverse orientation, generating sets of pGL3-3′(Target site)-FWD and pGL3-3′(Target site)-REV (Fig. 1A, bottom two plasmids). In these constructs, the sequences surrounding each individual target sequence were unchanged.

To analyze the effects of the ribosomes or associated helicases on the activities of siRNAs, we deleted the stop codons before the target site of pGL3-3′TAR-FWD, generating pGL3-3′TAR-FWD(–). In the presence of a stop codon, the ribosome complex would be released before reaching the target site while, in the absence of a stop codon, the target site in the TAR motif might be unwound and scanned by the ribosome (39–42).

In other cases, we chose genes for two similar firefly *Photinus pyralis* luciferases, namely, GL2 and GL3, whose nucleotide sequences are 95% homologous but whose computer-predicted secondary structures are quite different. Using these genes, we were able to target the same sequence within two different secondary structures and to examine the effects of the surrounding structure (pTAR-GL2-pGL3).

Finally, to examine the effects of the sequence and position of the target site within a noncoding region on RNAi, we placed the 3′-untranslated region of pTAR-GL2 in the forward (pGL3-3′Stop-FWD) and in the reverse (pGL3-3′Stop-REV) orientation, as well as within the 5′-untranslated region (pGL3-5′Stop-FWD and pGL3-5′Stop-REV; Fig. 1A, middle).

TAR is an unfavorable target for siRNAs regardless of its position

We examined the effects of target structure on RNAi using tightly structured TAR RNA as target. We varied the position of TAR and investigated RNAi both in the presence and in the absence of the Tat protein (Fig. 2A, top). We synthesized chemically five sets of siRNAs directed against the TAR sequence (Fig. 1B). Each of the five target sites (TAR1–TAR5) in the mRNA transcribed from pTAR-GL2 was strongly protected from cleavage by the siRNAs (Fig. 2A, top, black bars). Under the conditions of our experiments, the maximum RNAi effect was 50% at the TAR4 site. To examine positional effects, we placed the TAR motif downstream of the stop codon, generating pGL3-3′TAR-FWD, and compared the activities of several siRNAs. As indicated by the black and gray bars in the top panel in Figure 2A, siRNAs targeted to the TAR sequence were not very effective, irrespective of whether the target site was located before the initiation codon (black) or after the stop codon (gray). Furthermore, the patterns of efficacy were similar. These results support the hypothesis that efficiency of RNAi is more dependent on the structure and/or sequence of the target than on its position.

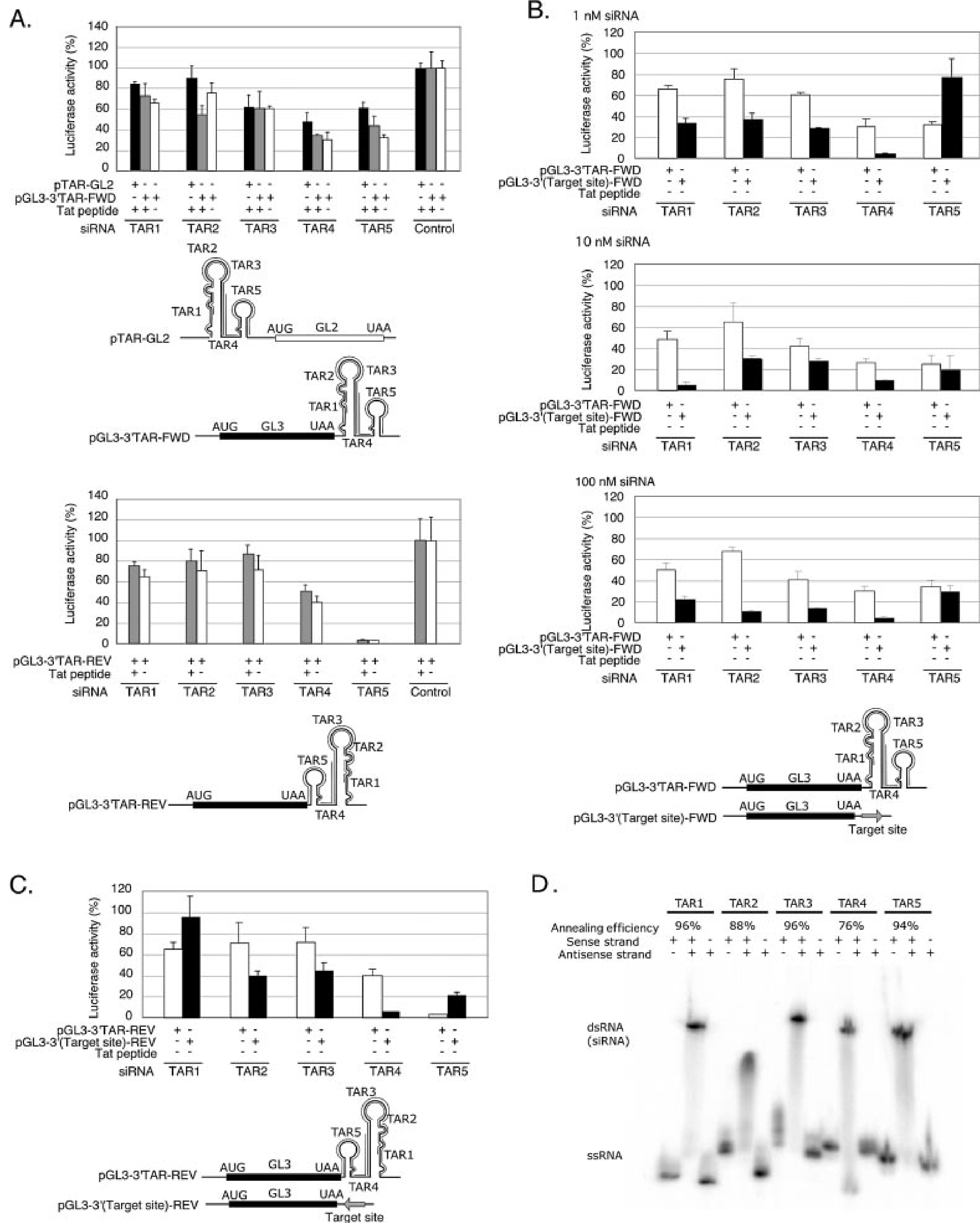

Figure 2. Inhibition of firefly luciferase activities by siRNAs targeted to TAR. (**A**) The inhibitory effects on pTAR-GL2, pGL3-3′TAR-FWD and pGL3-3′TAR-REV. Black bars show the remaining firefly luciferase activities when pTAR-GL2 was expressed in the presence of Tat protein. Gray and white bars show activities when pGL3-3′TAR-FWD (top) and pGL3-3′TAR-REV (bottom) were expressed in the presence and in the absence of Tat protein, respectively. (**B**) The inhibitory effects of 1, 10 and 100 nM siRNAs on pGL3-3′TAR-FWD (white bars) and pGL3-3′(Target site)-FWD (black bars) in the absence of Tat protein. (**C**) The inhibitory effects of siRNAs on pGL3-3′TAR-REV (white bars) and pGL3-3′(Target site)-REV (black bars) in the absence of Tat protein. (**D**) Annealing experiment for TAR1–TAR5 siRNAs using 5′ ^{32}P-labeled RNAs. Annealed siRNAs were separated from single-stranded RNA (ssRNA) by electrophoresis on a 20% native polyacrylamide gel.

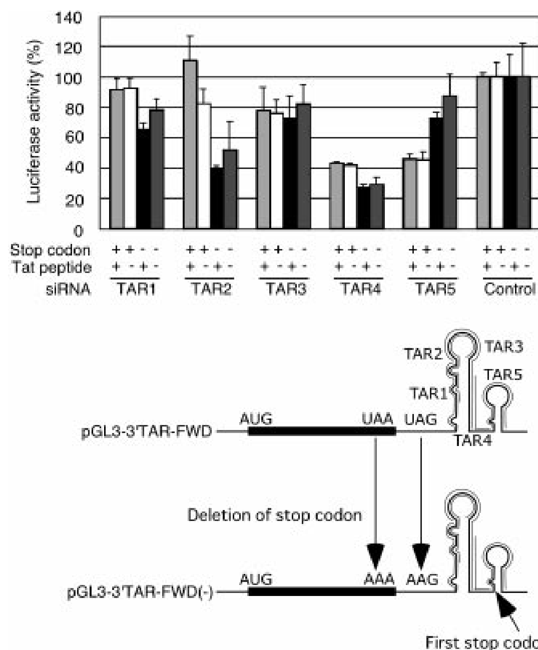

Figure 3. The effect of a stop codon before the target site of an siRNA when pGL3-3′TAR-FWD and pGL3-3′TAR-FWD(–) were expressed in the presence and in the absence of Tat protein. Gray and white bars show activities when pGL3-3′TAR-FWD (with stop codon) was expressed in the presence of and in the absence of Tat protein. Black and dark gray bars show activities when pGL3-3′TAR-FWD(–) was expressed in the presence and in the absence of Tat protein.

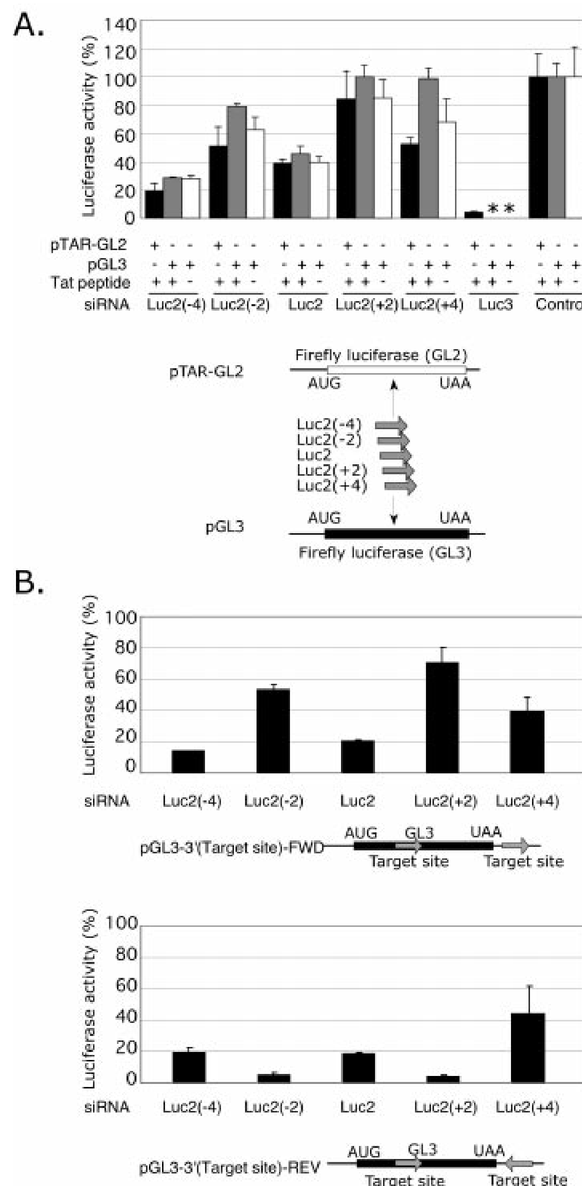

Figure 4. Inhibition of firefly luciferase activities by siRNAs targeted to sites around the Luc2 site. (**A**) The inhibitory effects on pTAR-GL2 and pGL3. Black bars show the remaining firefly luciferase activities when pTAR-GL2 was expressed in the presence of Tat protein. Gray and white bars show luciferase activities when pGL3 was expressed in the presence and in the absence of Tat protein, respectively. An asterisk indicates that the transfected siRNA did not have a sequence that allowed it to target the reporter mRNA. (**B**) The inhibitory effects of siRNAs on pGL3-3′(Target site)-FWD (top) and pGL3-3′(Target site)-REV (bottom) in the absence of Tat protein. Note that each siRNA had two target sites in each mRNA because the original plasmid pGL3 had the same target sequence around the Luc2 site as that in pTAR-GL2.

The Tat protein does not influence the effects of siRNAs directed against the TAR motif

To examine the effects on RNAi of the binding of the Tat protein to the TAR motif, we compared the efficiencies of siRNAs directed against the TAR motif in the presence and in the absence of a Tat expression vector. In the absence of the Tat expression vector, the absolute luciferase activity due to pTAR-GL2 was reduced dramatically (the reduction was at least 1000-fold, as anticipated from the putative Tat–TAR interaction and the Tat-mediated activation of transcription). Thus, in the absence of Tat protein, accurate and reproducible measurements were not possible (data not shown). Therefore, we made the comparison using pGL3-3′TAR-FWD (in which the TAR motif was at the 3′ end of the firefly luciferase gene). In this case, the Tat-mediated activation of pGL3-3′TAR-FWD was marginal since the gene was under the control of an SV40 promoter. Comparison of the effects of siRNAs revealed that the activities of the siRNAs showed a similar trend in the presence and in the absence of the Tat protein (Fig. 2A, top, gray and white bars), indicating that the effects of the binding of the Tat protein were marginal.

The Tat protein had, of course, no nonspecific effect on the activities of siRNAs, as demonstrated by the use of the reversed TAR sequence in pGL3-3′TAR-REV. In this case, the Tat protein was unable to bind to target sites (Fig. 2A, bottom, gray and white bars). The same was true for other target sites that lacked the TAR motif (around Luc2 site; see Fig. 4A, gray and white bars). Taken together, our results indicated that the Tat protein had no effects on the activities of the siRNAs, at least when the target sites were hidden within a tightly structured RNA.

The activities of siRNAs directed against the TAR motif are mainly affected by the structural environment

The effects of the siRNAs directed against the TAR motif were very small but it was unclear whether this phenomenon

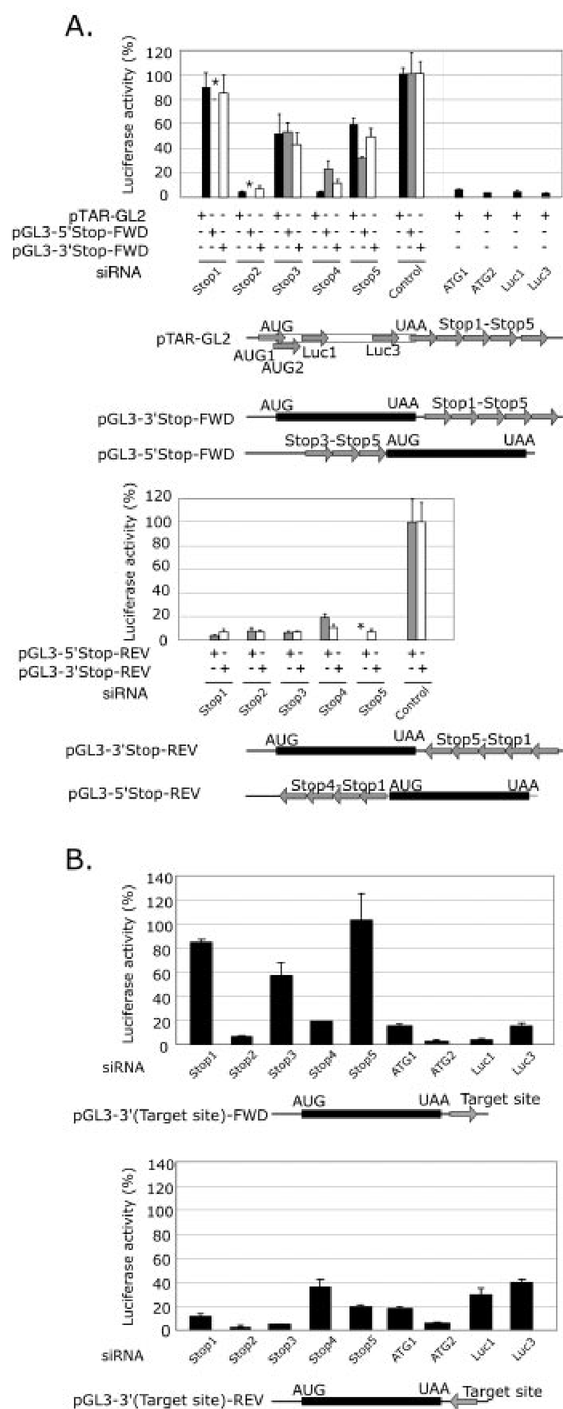

Figure 5. Inhibition of firefly luciferase activities by siRNAs targeted to a sequence that originated from the sequence after the stop codon of pTAR-GL2. (**A**) The inhibitory effects of siRNAs on the expression of pTAR-GL2, pGL3-5′Stop-FWD, pGL3-5′Stop-REV, pGL3-3′Stop-FWD and pGL3-3′Stop-REV in the presence of Tat protein. Black bars show the remaining firefly luciferase activities when pTAR-GL2 was expressed. Gray and white bars show those when pGL3-5′Stop-FWD or pGL3-5′Stop-REV and pGL3-3′Stop-FWD or pGL3-3′Stop-REV were expressed (top, results for FWD; bottom, results for REV), respectively. An asterisk indicates that the transfected siRNA did not have a sequence that allowed it to target the reporter mRNA. (**B**) The inhibitory effects on pGL3-3′(Target site)-FWD (top) and pGL3-3′(Target site)-REV (bottom) in the absence of the Tat protein.

Table 1. The siRNAs used in this study

Name	Sense and antisense sequences
TAR1	CUCUCUGGUUAGACCAGAU-CT AUCUGGUCUAACCAGAGAG-AC
TAR2	AGACCAGAUCUGAGCCUGG-GA CCAGGCUCAGAUCUGGUCU-AA
TAR3	GGGAGCUCUCUGGCUAACU-AG AGUUAGCCAGAGAGCUCCC-AG
TAR4	UGGCUAACUAGGGAACCCA-CT UGGGUUCCCUAGUUAGCCA-GA
TAR5	CUUAAGCCUCAAUAAAGCU-TG AGCUUUAUUGAGGCUUAAG-CA
ATG1	CGGUACUGUUGGUAAAAUG-GA CAUUUUACCAACAGUACCG-GA
ATG2	AUGGAAGACGCCAAAAACA-TA UGUUUUUGGCGUCUUCCAU-TT
Luc1	ACAUCACGUACGCGGAAUA-CT UAUUCCGCGUACGUGAUGU-TC
Luc2	GCUAUGAAACGAUAUGGGC-TG GCCCAUAUCGUUUCAUAGC-TT
Luc2(–4)	AGAAGCUAUGAAACGAUAU-GG AUAUCGUUUCAUAGCUUCU-GC
Luc2(–2)	AAGCUAUGAAACGAUAUGG-GC CCAUAUCGUUUCAUAGCUU-CT
Luc2(+2)	UAUGAAACGAUAUGGGCUG-AA CAGCCCAUAUCGUUUCAUA-GC
Luc2(+4)	UGAAACGAUAUGGGCUGAA-TA UUCAGCCCAUAUCGUUUCA-TA
Luc3	GUGCGUUGCUAGUACCAAC-CC GUUGGUACUAGCAACGCAC-TT
Stop1	UAAAAUGUAACUGUAUUCA-GC UGAAUACAGUUACAUUUUA-CA
Stop2	UGACGAAAUUCUUAGCUAU-TG AUAGCUAAGAAUUUCGUCA-TC
Stop3	AUACUCUAGAGGAUCUUUG-TG CAAAGAUCCUCUAGAGUAU-TA
Stop4	GGAACCUUACUUCUGUGGU-GT ACCACAGAAGUAAGGUUCC-TT
Stop5	CAUAAUUGGACAAACUACC-TA GGUAGUUUGUCCAAUUAUG-TC
GFP	GACGUAAACGGCCACAAGU-TT ACUUGUGGCCGUUUACGUC-TT

was due solely to the effects of structure. To clarify this issue, we constructed derivatives of pGL3 that included only a target site, without the rest of the TAR sequence, with each respective target sequence inserted at a common site, surrounded by the identical sequences. These derivatives were designated pGL3-3′(Target site)-FWD and pGL3-3′(Target site)-REV (Fig. 1A, bottom). The vectors had the same sequence apart from an inserted 23 bp sequence and none of the target sites was able to adopt a structure similar to its original structure in pTAR-GL2. In this way, we were able to evaluate the intrinsic activity associated with each target sequence.

We could not exclude the possibility that some or all of the inserted target sequences might have been able to form new base pairs with the surrounding sequences. However, in the case of target sites (Stop1–Stop5) which were not in a strict stem–loop structure, the efficiencies of the siRNAs were, in general, almost the same as those observed with the original reporter vector (pTAR-GL2), indicating that the effect of isolation of each target site was very small (see below and Fig. 5). In contrast, in the case of siRNAs targeted to the TAR

motif (Fig. 2B, top) and the reversed TAR motif (Fig. 2C), activities were higher at practically all positions when the target sequence was isolated from the tight hairpin-like structure (TAR1–TAR4). The same improvement in siRNA activity was observed even at higher concentrations of siRNA (10 or 100 nM; Fig. 2B, middle and bottom) indicating that the low efficiencies for structured target sites were not due to our choice of the low concentration of siRNAs (1 nM) used in this study.

Our results indicated that the low efficiencies of siRNAs directed against sequences around the TAR sequence were due to the tight structure itself, perhaps not exclusively but, at least, to a significant extent. Thus, in almost all cases (the exception being the less-structured TAR5 at 1 nM siRNA; Fig. 2B, top), the efficiency of silencing was significantly improved when the 23mer target regions were isolated from the highly structured TAR environment, indicating that highly structured local regions might interfere with the activity of RISC.

The possible participation of ribosomal complexes in the effects of siRNA on tight structures

Since the effects of siRNA depended to a significant extent on tight structure, we examined whether there were any differences in the effects of siRNAs when a stop codon was located upstream or downstream of a target site. We postulated that the siRNA/RISC might potentially associate with the ribosome or some other aspect of the translation machinery (43) that might unwind a structured mRNA (39–42). To examine the effects of ribosomes and/or associated helicases on the activities of siRNAs, we deleted the stop codons upstream of the target sites of pGL3-3′TAR-FWD to generate pGL3-3′TAR-FWD(–). We compared the efficiencies of siRNAs when a stop codon was located upstream and downstream of the target site. Using 1 nM siRNA (Fig. 3), we observed marginal enhancement of the activities of siRNAs when the stop codon was eliminated. However, the extent of the enhancement was significantly smaller than when the target site within the tight TAR motif was isolated in pGL3-3′(Target site)-FWD (Fig. 2B, top). Although we cannot rule out the participation of ribosome complexes in the mechanism of RNAi, it seems that their effects as a helicase might be small, at least in the case of the tight structure of TAR (44).

A small displacement of the target site changes the activities of siRNAs drastically but independently of the surrounding environment

Since our data suggested that structural effects were dominant when target sites were embedded in a tight RNA structure, while positional effects were marginal in such cases (Fig. 2A), we examined the effect of sequence itself and of its composition by sliding each target site by a few nucleotides. To this end, we chose the sensitive Luc2 site from among the three target sites (Luc1–Luc3) within the GL2 luciferase gene that we had tested in the past (26,44). This site was not a particularly effective target for siRNA both in our previous study (26) and in the present study (see Luc2 in Fig. 4A and Luc1, Luc3 in Fig. 5A). We prepared siRNAs targeted to sequences that were shifted by 2 or 4 bp from the original Luc2 site [Luc2(n), where n = –4, –2, +2, +4]. In the case of the Luc2 site, we observed ~60% inhibition (Fig. 4A, black bars). A shift of only two bases dramatically reduced the effect of siRNA [Luc2(+2)].

Next, to examine the environmental effects, we compared the efficiencies of siRNAs directed against the same target site in different mRNAs. For this purpose, we chose genes for two similar firefly *P.pyralis* luciferases, GL2 and GL3, because they encode the same amino acid sequence but the homology between the two genes at the nucleotide level is only 95%. We were able, therefore, to choose five identical target sequences around Luc2 in pTAR-GL2 and pGL3. Although the energetically most stable computer-predicted M-fold secondary structures of these mRNAs (45,46) are quite different from each other, the patterns of effects of siRNAs were surprisingly similar for the two constructs pTAR-GL2 and pGL3 (Fig. 4A, black and gray bars). These observations suggest that the effects of the siRNA might be governed by the target site and, more specifically, that the effects might be governed by the sequence itself without any effect of secondary structure, in the absence of a tight structure such as that adopted by the TAR motif.

Although an earlier analysis indicated a slight correlation between the activities of siRNAs and their GC content (30), the GC content of each of our siRNAs fell within a narrow range (32–47%). However, the effects of siRNA on Luc2 (47% GC) and on Luc2(+2) (42% GC) were, for example, very different. More recent studies, including our own, indicate that A or U at the 5′ end of the antisense strand significantly enhances activity of an siRNA (35,36,47,48). However, the relatively strong effect of Luc2 with a 5′-GCCC-3′ sequence having a tight 5′ end of the antisense strand cannot be explained by this relationship. Nevertheless, other results, including those shown in Figure 4B with forward and reversed targets, can be explained by the contribution of the A or U at the 5′ end of the antisense strand (see below).

The activities of siRNAs against sense and antisense target sequences are different

The sense target sites and the corresponding antisense target sites within the tight structure of the TAR motif were very similar with respect to the relative effects of siRNA (Fig. 2A). Therefore, we next examined the effects of siRNA on sense and antisense targets when the targets were not embedded in a tight RNA structure. We constructed derivatives of pGL3 with some targets of siRNAs in the forward orientation and some in the reverse orientation (Fig. 5A). We chose five target sites at, around and downstream of the stop codon in pTAR-GL2, and we placed this region either at a 5′ site (pGL3-5′Stop-FWD and pGL3-5′Stop-REV) or at a 3′ site (pGL3-3′Stop-FWD and pGL3-3′Stop-REV) in pGL3. When this region was placed at the 5′ site (pGL3-5′Stop-FWD), we included only three target sites (Stop3, Stop4 and Stop5) and omitted two other sites (Stop1 and Stop2) because the latter two sites produced an initiation codon, namely, AUG. Similarly, in the case of pGL3-5′Stop-REV, Stop5 was omitted because the reverse sequence of Stop5 created the initiation codon AUG.

Supporting the results obtained with the 5′ and 3′ TAR constructs, there were no positional effects when the target region was placed at the 5′ site or the 3′ site (compare results indicated by gray and white bars in Fig. 5A). As reported previously (49), significant effects of siRNA were clearly

detectable even when noncoding regions were targeted. Moreover, the effects of siRNAs on the reversed sequence were remarkably high (Fig. 5A, bottom). These results were unexpected in view of the results of the experiments with the TAR motif (Fig. 2A). We observed a similar discrepancy between the results for sense and antisense targets with other constructs that had a common surrounding environment (Figs 4B and 5B). The difference in siRNA efficiency between sense and antisense targets was also reported recently by Schwarz *et al*. (35).

Since the efficacy of our siRNAs depended mainly on the target sequence itself, we analyzed sequence preferences. We found, for example, that siRNAs with an A residue at the 19th nucleotide position from the 5′ end of the sense strand tended to have relatively high suppressive activities (TAR4 siRNA for the forward target; and TAR4, Stop1 and Stop2 siRNAs for the reverse target were effective). Moreover, siRNAs with a G residue at the 19th nucleotide in the sense strand tended to be less effective [Luc2(–2), Luc2(+2) and Stop3 siRNAs for the forward target and TAR1 siRNAs for the reverse target were ineffective]. Statistical analysis, based on our accumulated data, indicated that some nucleotides at specific positions are positively or negatively correlated with the efficiencies of siRNAs (Table 2). It is noteworthy that a similar preference (A19 in siRNA; U1 in the miRNA) was observed for miRNA sequences (50). This preference suggests a possible functional contribution of a U at the 5′ end of an antisense strand to the activities of both siRNA and miRNA. Recently, the importance of the low internal stability of the 5′ terminus of the antisense strand was also reported by Zamore's and Khvorova and Jayasena's groups (35,36). Furthermore, a U residue at 10th position in the sense strand (the middle nucleotide of the target) tends to be effective [TAR1 siRNA for the forward target and Luc2(–2), Luc2(+2), Stop1, Stop3 and Stop5 siRNAs for the reverse target were effective]. In our present experiments, these preferences can, by themselves, explain why reversed targets were attacked more effectively than the corresponding forward targets. We also found a significant negative correlation between the GC content of the 3′ half of siRNAs (in particular, from the 12th to the 19th nucleotide) and the activities of siRNAs. These tendencies can also be seen in another report by Vickers *et al.* (51).

Table 2. The influence of the 19th and 10th nucleotides of the target site[a] on siRNA activity

Nucleotide (sample no.)	siRNA activity >80%	>70%	>60%
Any nucleotide (162)	43 (26.5%)	66 (40.7%)	83 (51.2%)
A19 (51)	22 (43.1%)	26 (51.0%)	34 (66.7%)
U19 (51)	15 (29.4%)	23 (45.1%)	29 (56.9%)
G19 (42)	5 (11.9%)	15 (35.7%)	16 (38.1%)
C19 (18)	1 (5.6%)	2 (11.1%)	4 (22.2%)
U10 (52)	20 (38.5%)	24 (46.2%)	27 (51.9%)

[a]The siRNAs were designed to target the *Renilla* luciferase gene and the firefly luciferase gene (GL3). The 5′ nucleotides of the target sequences were located, sequentially, at positions 372–480 in the *Renilla* luciferase-coding sequence; and at positions 740–743, 747, 749, 751–756, 758–761, 763–765, 767, 769–791, 793, 795, 796, 798 and 800–805 in the firefly luciferase-coding sequence.

CONCLUDING REMARKS

To identify the major parameters that govern the effects of siRNA, we selected target sites for siRNAs in a coding region, in 5′- and 3′-untranslated regions, and in forward and reversed orientations. Our quantitative and systematic analysis of close to 50 different target sites revealed that the efficacy of siRNA was reduced when the target site was embedded within a tight RNA structure. Moreover, when a tight structure, such as that of TAR was involved, the effect of a protein, Tat, that interacts with the target was not significant. Positional effects also seemed unimportant, and the efficacy of siRNA appeared mostly to depend on the target sequence itself, with surrounding sequences having no major effects.

Selection of effective target sites is very important for the successful application of siRNA technology. In our experience, the number of effective target sites for siRNAs appears larger than that for conventional antisense molecules, such as ribozymes, perhaps because RISC has RNA helicase activity. Nevertheless, even with siRNAs, if the appropriate targets are not selected, an increase in dose fails to compensate for the ineffectiveness of the siRNAs. Therefore, as in the case of both antisense technology and ribozyme technology, the selection of the target site remains one of the most important determinants of success. In the present study, we demonstrated that it is the target sequence itself that is the major determinant of the effectiveness of an siRNA, and it is now possible to identify some of the preferences for nucleotides at specific positions (35–37,47,48; http://www.igene-therapeutics.co.jp). It remains to be determined whether RISC prefers such motifs for formation of an effective complex but, nonetheless, the information that we have obtained should be very useful for future selection of target sites and the successful application of siRNA technology.

REFERENCES

1. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
2. Fire,A. (1999) RNA-triggered gene silencing. *Trends Genet.*, **15**, 358–363.
3. Hammond,S.M., Caudy,A.A. and Hannon,G.J. (2001) Post-transcriptional gene silencing by double-stranded RNA. *Nature Rev. Genet.*, **2**, 110–119.
4. Sharp,P.A. (2001) RNA interference—2001. *Genes Dev.*, **15**, 485–490.
5. Zamore, P.D. (2001) RNA interference: listening to the sound of silence. *Nature Struct. Biol.*, **8**, 746–750.
6. McManus,M.T. and Sharp,P.A. (2002) Gene silencing in mammals by small interfering RNAs. *Nature Rev. Genet.*, **3**, 737–747.
7. Akashi,H., Miyagishi,M. and Taira,K. (2002) Suppression of gene expression by RNA interference in cultured plant cells. *Antisense Nucleic Acid Drug Dev.*, **11**, 359–367.
8. Dykxhoorn,D.M., Novina,C.D. and Sharp,P.A. (2003) Killing the messenger: short RNAs that silence gene expression. *Nature Rev. Mol. Cell Biol.*, **4**, 457–467.
9. Bernstein,E., Caudy,A.A., Hammond,S.M. and Hannon,G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363–366.
10. Provost,P., Dishart,D., Doucet,J., Frendewey,D., Samuelsson,B. and Rådmark,O. (2002) Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J.*, **21**, 5864–5874.
11. Zhang,H., Kolb,F.A., Brondani,V., Billy,E. and Filipowicz,W. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.*, **21**, 5875–5885.

12. Kawasaki,H., Suyama,E., Iyo,M. and Taira,K. (2003) siRNAs generated by recombinant human Dicer induce specific and significant but target site-independent gene silencing in human cells. *Nucleic Acids Res.*, **31**, 981–987.
13. Song,J.J., Liu,J., Tolia,N.H., Schneiderman,J., Smith.S.K., Martienssen,R.A., Hannon,G.J. and Joshua-Tor,L. (2003) The crystal structure of the Argonaute2 PAZ domain reveals an RNA-binding motif in RNAi effector complexes. *Nature Struct. Biol.*, **10**, 1026–1032.
14. Lingel,A., Simon,B., Izaurralde,E. and Sattler,M. (2003) Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature*, **426**, 465–469.
15. Yan,K.S., Yan,S., Farooq,A., Han,A., Zeng,L. and Zhou,M.M. (2003) Structure and conserved RNA binding of the PAZ domain. *Nature*, **426**, 469–474.
16. Hammond,S.M., Bernstein,E., Beach,D. and Hannon,G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**, 293–296.
17. Martinez,J., Patkaniowska,A., Urlaub,H., Lührmann,R. and Tuschl,T. (2000) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*, **110**, 563–574.
18. Hammond,S.M., Boettcher,S., Caudy,A.A., Kobayashi,R. and Hannon,G.J. (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science*, **293**, 1146–1150.
19. Nykänen,A., Haley,B. and Zamore,P.D. (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell*, **107**, 309–321.
20. Fraser,A.G., Kamath,R.S., Zipperlen,P., Martinez-Campos,M., Sohrmann,M. and Ahringer,J. (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature*, **408**, 325–330.
21. Gönczy,P., Echeverri,G., Oegema,K., Coulson,A., Jones,S.J., Copley,R.R., Duperon,J., Oegema,J., Brehm,M., Cassin,E. *et al.* (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature*, **408**, 331–336.
22. Elbashir,S.M., Lendeckel,W. and Tuschl,T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.*, **15**, 188–200.
23. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.
24. Caplen,N.J., Parrish,S., Imai,F., Fire,A. and Morgan,R.A. (2001) Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl Acad. Sci. USA*, **98**, 9742–9747.
25. Tuschl,T. (2002) Expanding small RNA interference. *Nat. Biotechnol.*, **20**, 446–448.
26. Miyagishi,M. and Taira,K. (2002) U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.*, **20**, 497–500.
27. Lee,N.S., Dohjima,T., Bauer,G., Li,H., Li,M.J., Ehsani,A., Salvaterra,P. and Rossi,J. (2002) Expression of small interfering RNAs targeted against HIV-1 *rev* transcripts in human cells. *Nat. Biotechnol.*, **20**, 500–505.
28. Paul,C.P., Good,P.D., Winer,I. and Engelke,D.R. (2002) Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.*, **20**, 505–508.
29. Brummelkamp,T.R., Bernards,R. and Agami,R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science*, **296**, 550–553.
30. Yu,J.Y., DeRuiter,S.L. and Turner,D.L. (2002). RNA interference by expression of short interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 6047–6052.
31. Paddison,P.J., Caudy,A.A., Bernstein,E., Hannon,G.J. and Conklin,D.S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.*, **16**, 948–958.
32. Sui,G., Soohoo,C., Affar,el B., Gay,F., Shi,Y. and Forrester,W.C. (2002) A DNA vector based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 5515–5520.
33. Xia,H., Mao,Q., Paulson,H.L. and Davidson,B.L. (2002) siRNA-mediated gene silencing *in vitro* and *in vivo*. *Nat. Biotechnol.*, **20**, 1006–1010.
34. Holen,T., Amarzguioui,M., Wiiger,M.T., Babaie,E. and Prydz,H. (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res.*, **30**, 1757–1766.
35. Schwarz,D.S., Hutvágner,G., Du,T., Xu,Z., Aronin,N. and Zamore,P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.
36. Khvorova,A., Reynolds,A. and Jayasena,S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell*, **115**, 199–208.
37. Taira,K., Miyagishi,M. and Matsumoto,S. (2003) Enhancement of RNAi activity through mutations. US Patent Application.
38. Boutla,A., Delidakis,C., Livadaras,I., Tsagris,M. and Tabler,M. (2001) Short 5′-phosphorylated double-stranded RNAs induce RNA interference in *Drosophila*. *Curr. Biol.*, **11**, 1776–1780.
39. Kawasaki,H., Onuki,R., Suyama,E. and Taira,K. (2002) Identification of genes that function in the TNF-alpha-mediated apoptotic pathway using randomized hybrid ribozyme libraries. *Nat. Biotechnol.*, **20**, 376–380.
40. Kawasaki,H. and Taira,K. (2002) Identification of genes by hybrid ribozymes that couple cleavage activity with the unwinding activity of an endogenous RNA helicase. *EMBO Rep.*, **3**, 443–450.
41. Hanes,J., Jermutus,L., Schaffitzel,C. and Plückthun,A. (1999) Comparison of *Escherichia coli* and rabbit reticulocyte ribosome display systems. *FEBS Lett.*, **450**, 105–110.
42. Sawata,S. and Taira,K. Modified peptide selection *in vitro* by introduction of a protein–RNA interaction. *Protein Eng.*, in press.
43. Djikeng,A., Shi,H., Tschudi,C., Shen,S. and Ulli,E. (2003) An siRNA ribonucleoprotein is found associated with polyribosomes in *Trypanosoma brucei*. *RNA*, **9**, 802–808.
44. Warashina,M., Kuwabara,T., Kato,Y., Sano,M. and Taira,K. (2001) RNA–protein hybrid ribozymes that efficiently cleave any mRNA independently of the structure of the target RNA. *Proc. Natl Acad. Sci. USA*, **98**, 5572–5577.
45. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
46. Zuker,M., Mathews,D.H. and Turner,D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: A practical guide. In Barciszewski,J. and Clark,B.F.C. (eds), *RNA Biochemistry and Biotechnology*, Kluwer Academic Publishers, NATO ASI Series, pp. 11–43.
47. Miyagishi,M. and Taira,K. (2003) Strategies for generation of an siRNA expression library directed against the human genome. *Oligonucleotides*, **13**, 325–333.
48. Taira,K. and Miyagishi,M. (2003) Methods and devices for the prediction of favorable target sites of siRNA. Japanese Patent Application 2003-349283.
49. Yokota,T., Sakamoto,N., Enomoto,N., Tanabe,Y., Miyagishi,M., Maekawa,S., Yi,L., Kurosaki,M., Taira,K., Watanabe,M. and Mizusawa,H. (2003) Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs. *EMBO Rep.*, **4**, 602–608.
50. Reinhart,B.J. and Bartel,D.P. (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science*, **297**, 1831.
51. Vickers,T.A., Koo,S., Bennett,C.F., Crooke,S.T., Dean,N.M. and Baker,B.F. (2003) Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *J. Biol. Chem.*, **278**, 7108–7118.