DIABLO is essential for the anti-apoptotic activity of survivin during taxol-induced apoptosis. *J. Biol. Chem.* **278:** 23130–23140.
41. Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO J.* **20:** 6877–6888.
42. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* **13:** 3191–3197.
43. Hutvagner, G., and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297:** 2056–2060.
44. Semizarov, D., *et al.* (2003). Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl. Acad. Sci. USA* **100:** 6347–6352.
45. Jackson, A. L., *et al.* (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21:** 635–637.
46. Yoshida, Y., *et al.* (1998). Generation of fiber-mutant recombinant adenoviruses for gene therapy of malignant glioma. *Hum. Gene Ther.* **9:** 2503–2515.
47. Kanegae, Y., *et al.* (1995). Efficient gene activation in mammalian cells by using recombinant adenovirus expressing site-specific Cre recombinase. *Nucleic Acids Res.* **23:** 3816–3821.
48. Nyberg-Hoffman, C., Shabram, P., Li, W., Giroux, D., and Aguilar-Cordova, E. (1997). Sensitivity and reproducibility in adenoviral infectious titer determination. *Nat. Med.* **3:** 808–811.
49. Zhang, W. W., Koch, P. E., and Roth, J. A. (1995). Detection of wild-type contamination in a recombinant adenoviral preparation by PCR. *Biotechniques* **18:** 444–447.
50. Uchida, H., *et al.* (2003). 5-Fluouracil efficiently enhanced apoptosis induced by adenovirus-mediated transfer of caspase-8 in DLD-1 colon cancer cells. *J. Gene Med.* **5:** 287–299.

Published online 3 February 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

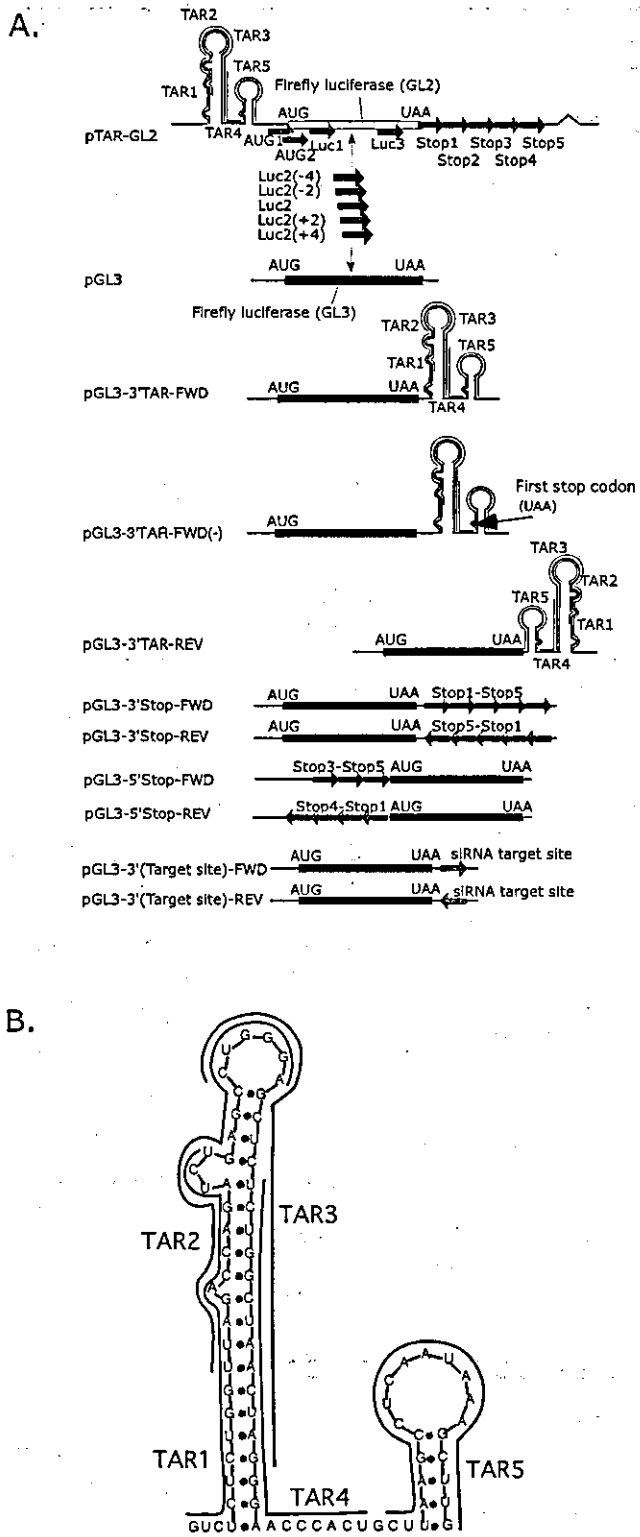

Figure 1. The reporter mRNAs for firefly luciferases that were used in the present study. (**A**) The white box represents the firefly luciferase gene GL2 and the black boxes represent the firefly luciferase gene GL3. TAR motifs are shown with their putative secondary structures at their respective positions 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 TGA ATT TCA CGG CGA TCT TTC C, and GGG AAG ATC GCC GTG AAA TTC AAG 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.

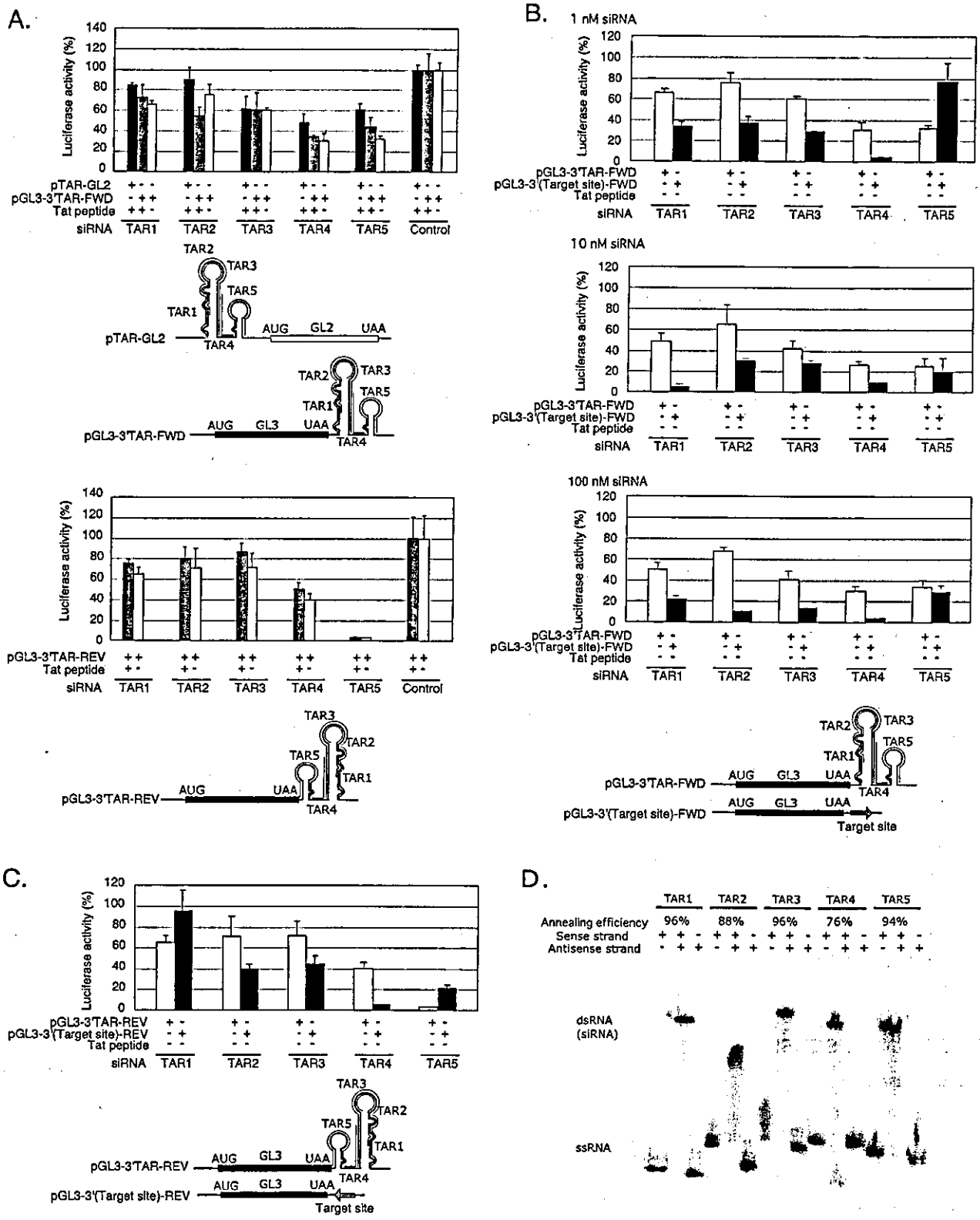

Figure 2. Inhibition of firefly luciferase activities by siRNAs targeted to TAR. (**A**) The inhibitory effects on pTAR-GL2, pGL3-3′TAR-FWD and pGL3-3′TAR-REV. Black bars show the remaining firefly luciferase activities when pTAR-GL2 was expressed in the presence of Tat protein. Gray and white bars show 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.

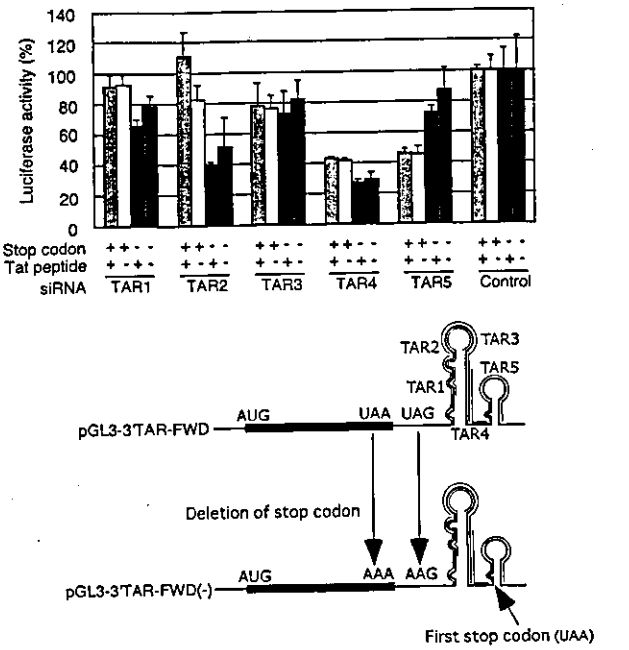

Figure 3. The effect of a stop codon before the target site of an siRNA when pGL3-3′TAR-FWD and pGL3-3′TAR-FWD(–) were expressed in the presence and in the absence of Tat protein. Gray and white bars show activities when pGL3-3′TAR-FWD (with stop codon) was expressed in the presence of and in the absence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

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

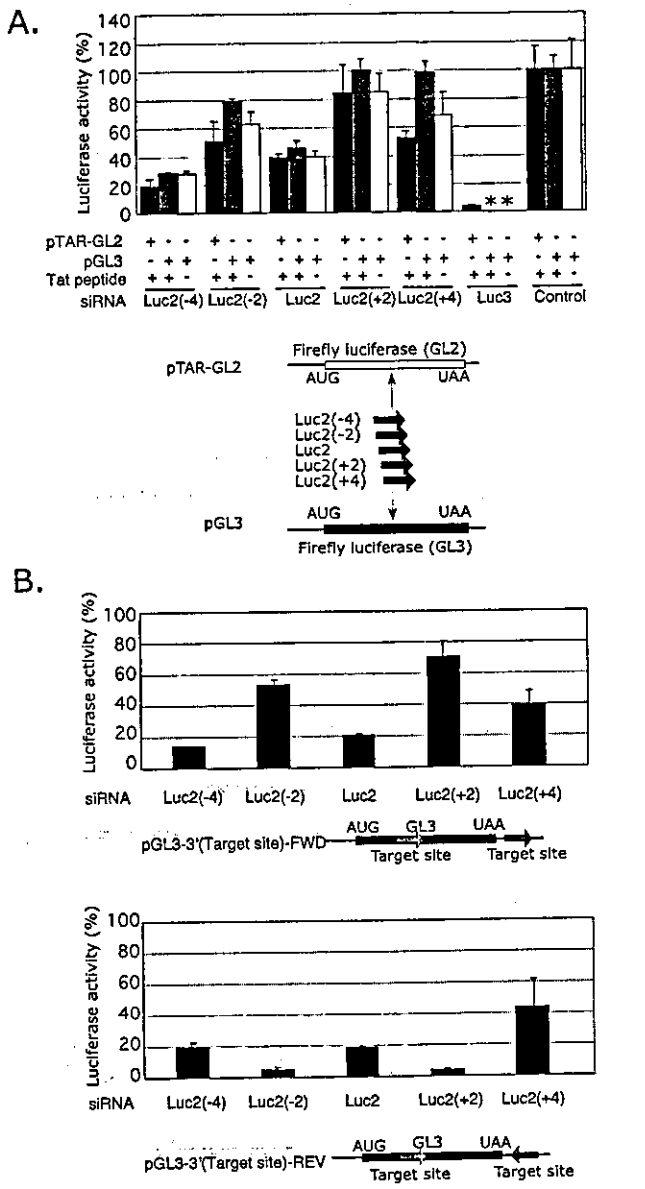

Figure 4. Inhibition of firefly luciferase activities by siRNAs targeted to sites around the Luc2 site. (**A**) The inhibitory effects on pTAR-GL2 and pGL3. Black bars show the remaining firefly luciferase activities when pTAR-GL2 was expressed in the presence of Tat protein. Gray and white bars show 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 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

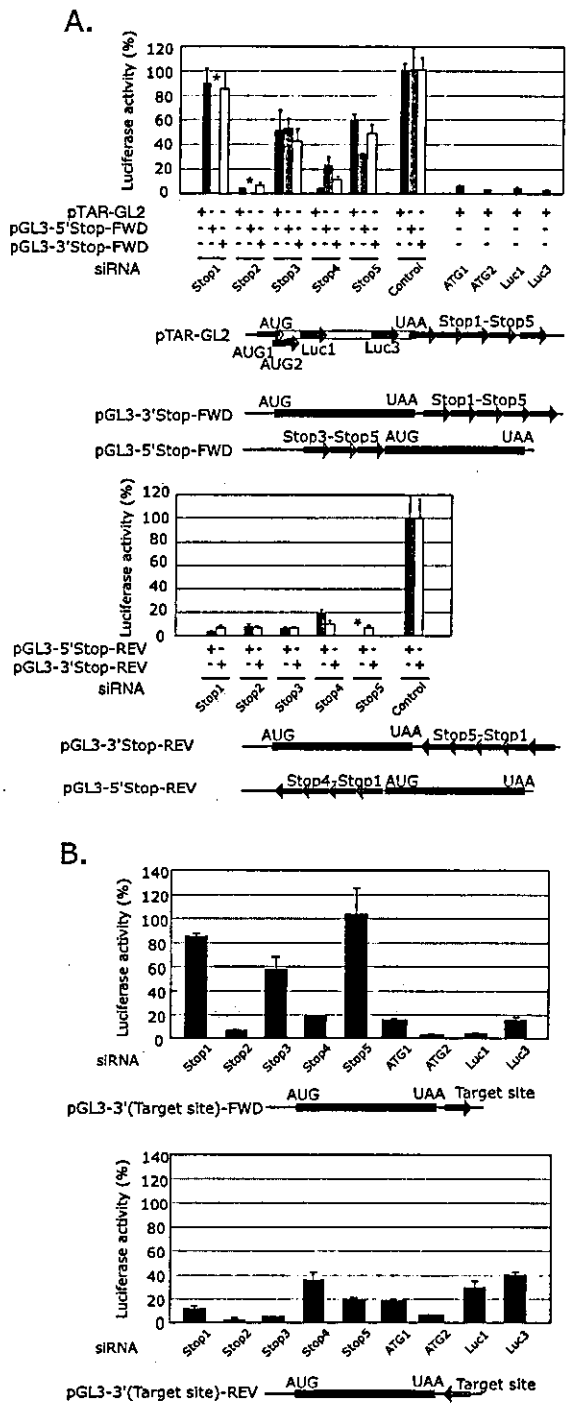

Figure 5. Inhibition of firefly luciferase activities by siRNAs targeted to a sequence that originated from the sequence after the stop codon of pTAR-GL2. (**A**) The inhibitory effects of siRNAs on the expression of pTAR-GL2, pGL3-5'Stop-FWD, pGL3-5'Stop-REV, pGL3-3'Stop-FWD and pGL3-3'Stop-REV in the 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.

OLIGONUCLEOTIDES 13:325–333 (2003)

Strategies for Generation of an siRNA Expression Library Directed Against the Human Genome

MAKOTO MIYAGISHI and KAZUNARI TAIRA

ABSTRACT

RNA interference (RNAi) is a phenomenon whereby expression of an individual gene is specifically silenced by the introduction of a double-stranded RNA (dsRNA) whose sequence is homologous to that of the gene in question. The generation of a small interfering RNA (siRNA) expression library directed against the entire human genome is a project that requires solutions to many difficult technical problems. We present here some strategies for solving some of these problems, including the development of genetically stable and highly active siRNA expression vectors, a procedure for selection of favorable target sites, and an efficient and inexpensive procedure for constructing an siRNA expression library.

INTRODUCTION

THE HUMAN GENOME includes many genes whose functions are unknown. Both hybrid ribozyme technology (Warashina et al., 2001; Kawasaki and Taira, 2002; Kawasaki et al., 2002; Suyama et al., 2003a,b) and RNA interference (RNAi) appear to provide methods that might reveal such functions (McManus and Sharp, 2003).

RNAi is an evolutionarily conserved phenomenon in plants and animals whereby double-stranded RNA (dsRNA) induces the sequence-specific degradation of homologous RNA (Fire et al., 1998; for review, see Zamore, 2001). RNAi operates effectively in various organisms, but the experimental silencing of specific genes by RNAi initially was difficult in mammalian systems because of the nonspecific dsRNA-dependent inhibition of protein synthesis via the PKR pathway and the nonspecific degradation of RNA that occurs on activation of RNase L (Elbashir et al., 2001). However, the recent work by Tuschl's group demonstrated that 21-nt or 22-nt RNAs with 2-nt 3′-overhangs (small interfering RNAs [siRNAs]) can silence the expression of target genes without nonspecific inhibition of gene expression in cultured mammalian cells (Caplen et al., 2001; Elbashir et al., 2001). Various groups, including our own, have developed systems for vector-mediated specific RNAi in mammalian cells (Brummelkamp et al., 2002; Kawasaki and Taira, 2003; Lee et al., 2002; Miyagishi and Taira, 2002; Paddison et al., 2002; Sui et al., 2002; Paul et al., 2002; Yu et al., 2002; for reviews, see McManus and Sharp, 2002; Tuschl, 2002). Each of these vector systems exploits a polymerase III (pol III) promoter, such as a U6, H1, or tRNA promoter, and the various systems can be divided into two groups depending on whether the expressed RNA is tandem type or hairpin type.

Gene silencing by siRNA is highly specific and is extremely effective. Thus, genomewide and comprehensive analysis of gene expression in various organisms should be facilitated by exploitation of this phenomenon. Recently, systematic functional analysis by RNAi of a number of genes was performed in *Caenorhabditis elegans*, and the functions of many genes were identified (Fraser et al., 2000; Gonczy et al., 2000). In the near future, a similar approach using a genomewide RNAi library should be possible in mammalian cells when siRNA oligonucleotide libraries or siRNA expression libraries be-

Department of Chemistry and Biotechnology, School of Engineering, University of Tokyo, Tokyo 113-8656, Japan, and Gene Function Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Science City 305-8562, Japan.

come available. Randomized ribozyme libraries have already been used for identification of novel genes (Beger et al., 2001; Kawasaki and Taira, 2002; Kawasaki et al., 2002; Kruger et al., 2000; Li et al., 2000; Onuki et al., 2002; Rhoades and Wong-Staal, 2003; Suyama et al., 2003a,b; Welch et al., 2000), and several groups, including our own, have begun to generate siRNA libraries directed against the entire human genome. In this report, we discuss the practical problems that are faced in efforts to construct an siRNA expression library and methods for solving some of them.

MATERIALS AND METHODS

Constructs for transient transfections

We prepared tandem type and hairpin type siRNA expression vectors using the piGENE hU6 vector (Miyagishi and Taira, 2003; iGENE Therapeutics, Inc.; *www.iGENE-therapeutics.co.jp*), which contains a human U6 promoter and two *Bsp*MI sites. For construction of tandem type siRNA expression plasmids, we amplified DNA fragments that included sequences of sense and antisense regions and the U6 promoter by PCR, using the piGENE hU6 vector as template and primers that included the sequences of the sense or antisense sequence plus the terminator. After digestion of products of PCR with *Bsp*MI, each fragment was ligated into the *Bsp*MI sites of the piGENE hU6 vector to yield a series of siRNA expression vectors. For construction of hairpin type siRNA expression vectors, we synthesized oligonucleotides with the hairpin sequence, the terminator sequence, and overhanging sequences. Then, we annealed the fragments and inserted them into the *Bsp*MI sites of the piGENE hU6 vector. H1 promoter-driven siRNA expression vectors were generated as described by Brummelkamp et al. (2002). Sequences inserted immediately downstream of the U6 promoters or the H1 promoters were as follows (only the sense sequences are shown): in pU6tandem21, pU6hairpin21, pH1hairpin21, and pH1S hairpin21, GTG CGC TGC TGG TGC CAA CCC. The sequences of the various mutated constructs are indicated in Figure 2.

Culture and transfection of cells and assays of expression of reporter genes

HeLa S3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Transfections were performed with Lipofectamine™ 2000 reagent (Life Technologies, Rockville, MD) using 30 ng of siRNA expression plasmid, 30 ng of a firefly luciferase expression plasmid (pGL3) (Promega, Madison, WI), and 10 ng of a *Renilla* luciferase expression plasmid (pRL-RSV) (Miyagishi et al., 2001), as described by Promega. Luciferase activities were analyzed after 24 hours with a Dual Luciferase System (Promega). To ensure that equal amounts of DNA were used in each transfection, empty plasmids were included at appropriate levels as necessary.

Algorithm for identification of favorable target sites

Predictions of the effects of the RNAi were made by the partial least-squares (PLS) method. We analyzed 114 sets of siRNA sequences and the corresponding suppression values, which we obtained from published data and our own experiments, and extracted significant correlated factors, including GC contents. We then prepared a PLS calibration model using the sequences of target sites, the corresponding suppression values, and the extracted factors using a PLS1 algorithm of our own making (Lindberg et al., 1983). The optimum number of PLS factors was determined by cross-validation, that is, the best calibration as judged by the lowest standard error of cross-validation. After equations had been established, suppression values for various sites in the gene for firefly luciferase were predicted using the respective target sequences.

RESULTS AND DISCUSSION

Optimization of siRNA expression systems: Tandem type or hairpin type?

In efforts to construct an siRNA expression library, it is first necessary to optimize the siRNA expression system to ensure that it has high suppressive activity and high genetic stability.

Most siRNA expression systems exploit a pol III promoter, such as the U6 or H1 promoter, and such systems can be divided into two groups, tandem type and hairpin type, depending on the RNA that is expressed (Fig. 1). In the case of tandem type siRNA expression vectors, both the sense and the antisense strands are driven separately by their own respective pol III promoters. They anneal inside cells and form siRNA duplexes with approximately 4-nt overhangs at each 3′-end (Lee et al., 2002; Miyagishi and Taira, 2002). In contrast, in the case of hairpin type siRNA expression vectors, sense and antisense oligonucleotides are connected by a loop and are expressed as a single macromolecule (Brummelkamp et al., 2002; Kawasaki and Taira, 2003; Paddison et al., 2002; Paul et al., 2002; Sui et al., 2002; Taira and Miyagishi, 2001; Yu et al., 2002). The transcribed RNA rapidly forms a hairpin structure with a stem and a loop and is, apparently, processed to siRNA by Dicer, as is miRNA, which is endogenous small non-coding RNA derived from stem-loop precursor. (Ketting et al., 2001).

To determine which vector system had the highest suppressive activity, we constructed both types of siRNA ex-

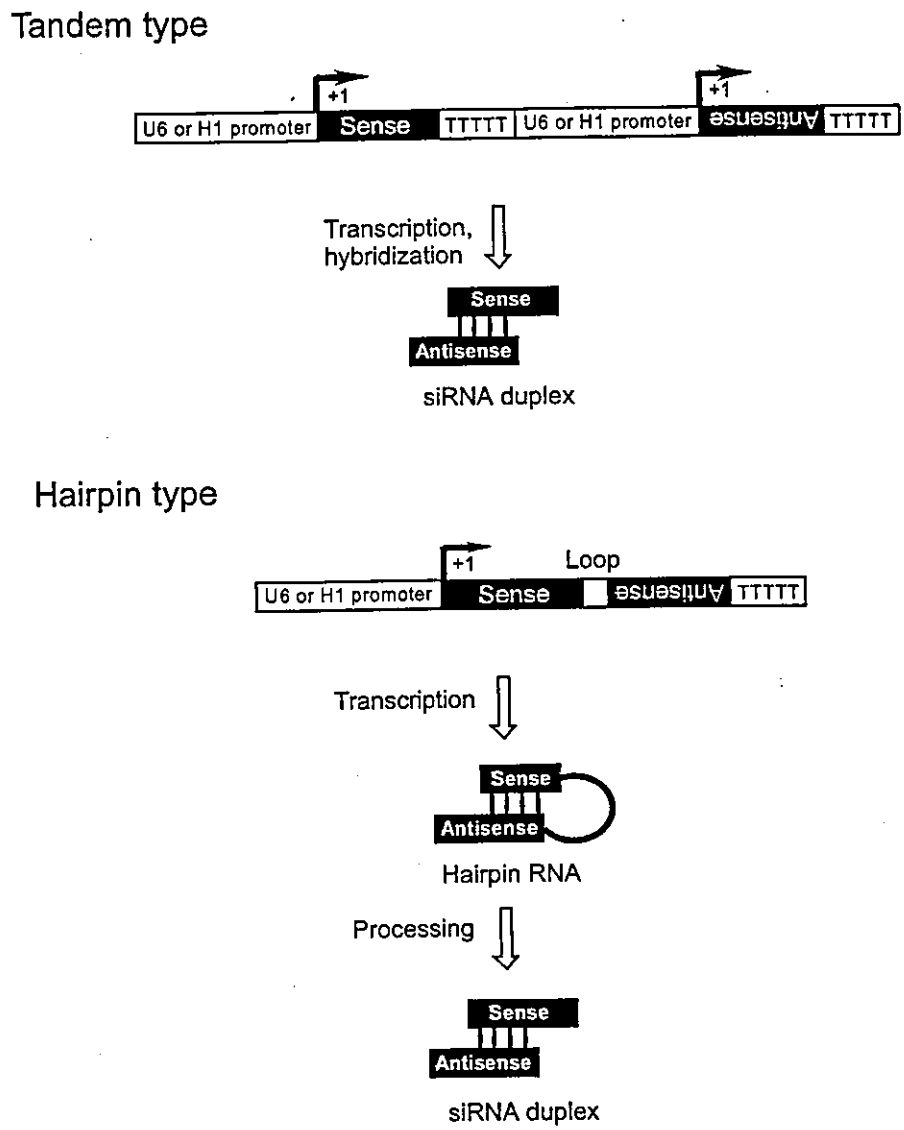

FIG. 1. Schematic representation of tandem type and hairpin type siRNA expression vectors.

pression vector targeted against the same site in the firefly gene for luciferase and compared their suppressive activities. Although both types were associated with high levels of suppressive activity when the concentration of siRNA expression vector was high, the hairpin type (stem length 21 nt) had significantly higher suppressive activity than the tandem type at low concentrations of the siRNA expression vector (Fig. 2A).

Genetically stable hairpin type siRNA expression vectors

The hairpin type siRNA expression vector had higher suppressive activity than the tandem type siRNA expression vector. However, the palindromic sequence of hairpin type siRNA expression vectors frequently is the cause of two serious technical problems. It is difficult to determine the sequences of constructs that contain a hairpin region, and, furthermore, from 20% to 40% of such constructs accumulate mutations within the hairpin region on introduction into *Escherichia coli*. This latter phenomenon presents the most serious problem in attempts to construct a library. Our studies had indicated that we could sequence constructs with two or more point mutations or insertions/deletions in the sense or antisense region without any problems and also that some of these constructs retained suppressive activity. Therefore, we examined the activities of constructs with various mutations and insertions/deletions in the sense or antisense region in detail (the complete results of this analysis will be published elsewhere). As shown in Figure 2A, as many as five point mutations from C to T, which generated

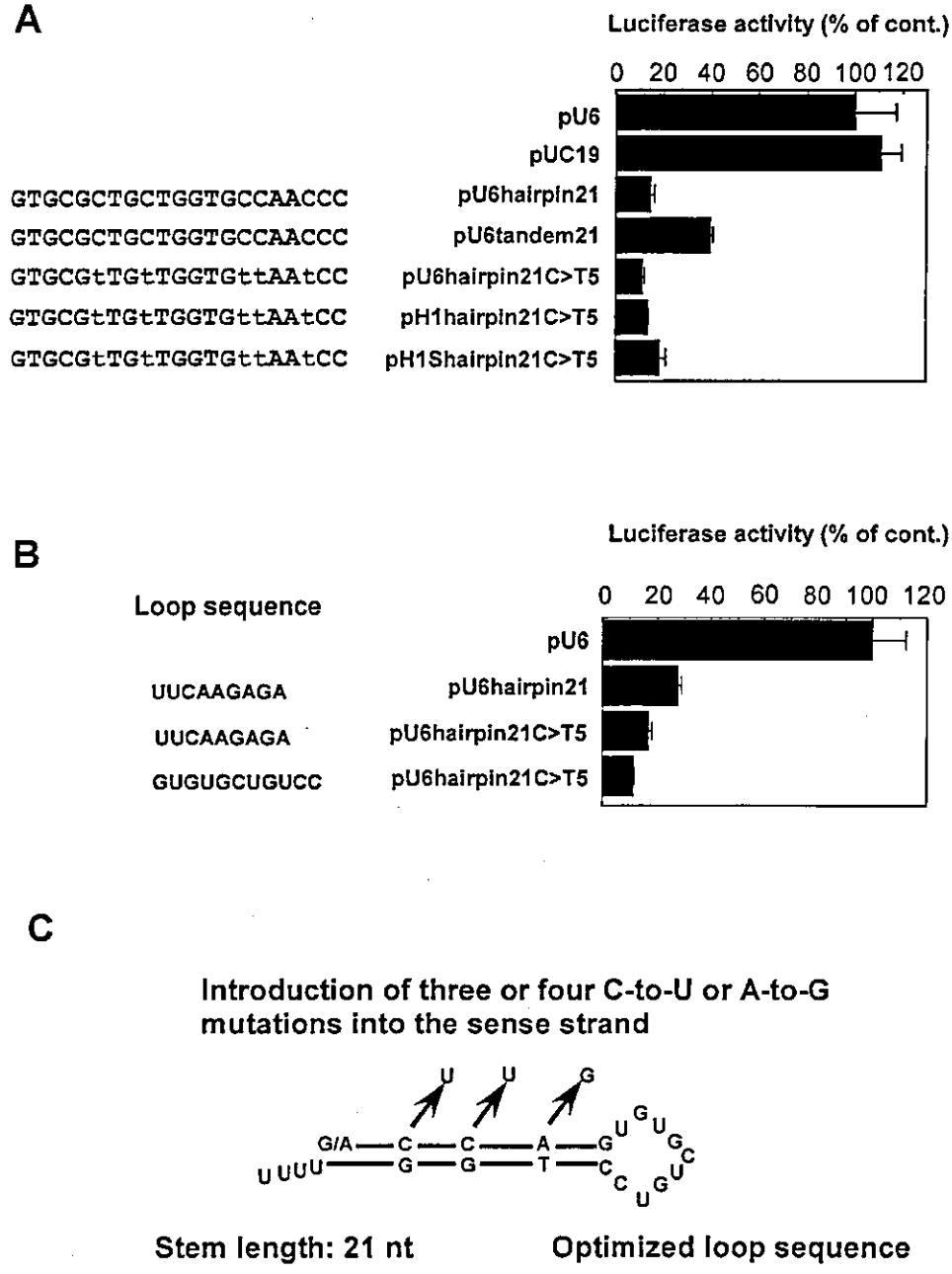

FIG. 2. (A) Comparative analysis of the effects of the U6 promoter-driven 21-nt tandem type and 2-nt hairpin type siRNA expression vectors, the H1 promoter-driven 21-nt hairpin type siRNA expression vector, and the U6 promoter-driven 21-nt hairpin type siRNA expression vector with various mutations in the hairpin region. HeLa S3 cells were cotransfected with 10 ng pRL-RSV, 30 ng pGL3, and 30 ng individual siRNA expression vectors targeted to the firefly gene for luciferase. The C>T5 and C>T1 represent constructs that have five and one C to T mutations in the sense strand, respectively. The activities of firefly luciferase were normalized by reference to those of *Renilla* luciferase. Each bar indicates an average value, and vertical bars indicate standard errors (SE) of triplicate assays. (B) The effects of the loop sequence on silencing. HeLa S3 cells were cotransfected with 10 ng pRL-RSV, 30 ng pGL3, and 30 ng siRNA expression vector with the indicated loop sequence, targeted to the firefly gene for luciferase. (C) Schematic representation of the proposed siRNA expression system.

G:U base-pairing, in the sense strand did not affect the silencing effect (Fig. 2A, pU6hairpin21C>T5).

While generating these constructs with mutated hairpin structures, we noted that the rates of mutation in *E. coli* of the hairpin region of the constructs were markedly reduced when we inserted more than three mutations in the sense or antisense strand (Taira and Miyagishi, 2001). The stability of such constructs is obviously important for the mainte-

nance of the corresponding plasmids in their prokaryotic host. Collectively, our data suggest that one way to generate a useful hairpin type siRNA expression vector for construction of an siRNA expression library is to introduce multiple C to T (or A to G) mutations into the sense strand of the hairpin exclusively (Fig. 2B).

Comparative analysis of effects of various pol III promoters on silencing activity

We compared the RNAi activity of siRNAs whose transcription was controlled by the U6 promoter and the H1 promoter, both of which are type II polymerase III promoters. These promoters contain the same consensus motifs, namely, a distal sequence element (DSE), a proximal sequence element (PSE), and a TATA box. As shown in Figure 2A, the H1 promoter-driven siRNA construct targeted against the firefly gene for luciferase had activity that was similar to but slightly lower than that of the U6-driven siRNA construct (Fig. 2A, compare pH1hairpin21C>T5, with promoter length of ~230 bp, or pH1S hairpin21C>T5, with promoter length of ~100 bp, with pU6hairpin21C>T5) (unpublished observations). A tRNA-linked siRNA expression vector was developed recently as an expression system that includes transport of hairpin RNA to the cytoplasm (Kawasaki and Taira, 2003), but the longer stem (30 nt) required makes such a system more expensive. Therefore, we chose a U6-driven siRNA expression vector for the construction of our siRNA expression library.

Optimization of loop sequence

A variety of loop sequences has been used by various researchers for expression of hairpin type siRNAs. For example, Agami's group used a 9-nt loop sequence, 5′-UUCAAGAGA-3′ (Brummelkamp et al., 2002), whereas Paul et al. (2002) used a 4-nt, 5′-UUCG-3′, sequence. In our experiments using a hairpin type siRNA expression vector with randomized 6-nt loop sequences, we found that hairpin RNAs with different loop sequences had different suppressive activities (Kawasaki and Taira, 2003a; data not shown). Therefore, the loop sequence appears to influence the RNAi effect of siRNAs.

We wondered if the natural loop sequences of miRNAs might be preferable to the loop of hairpin RNAs generated from hairpin type siRNA expression vectors. Examination of several miRNA-derived loop sequences revealed that loop 7, which is 11 nt long and is derived from human miR-26b (5′-GUG UGC UGU CC-3′), was extremely effective (the loop shown in Fig. 2B), and we adopted this loop as the optimized loop sequence in the hairpin type siRNA expression vector for construction of the siRNA expression library.

Selection of favorable target sites

One of the most important and critical problems associated with construction of an siRNA expression library is the selection of target sites. The effectiveness of siRNA is strongly dependent on the target site in the mRNA, and the probability of an siRNA directed against an arbitrarily selected target site having high suppressive activity seems, empirically, to be between approximately 20% and 40%. To generate a high-quality siRNA expression library, it is necessary to develop an algorithm that can predict sites likely to be effectively targeted. We analyzed datasets for the activities of siRNAs directed against several hundred target sites from published reports and our own experiments (unpublished observations). We extracted several correlated factors and generated an algorithm by a nonlinear regression method (details will be published elsewhere).

Figure 3 is a plot of predicted vs. actual results for siRNAs targeted against the firefly gene for luciferase, which were not used for generation of the algorithm. The correlation coefficient was not very high (0.58), but notably, all siRNAs with a predicted score (relative suppressive activity) of >0.73 had strong suppressive activity (>90% suppression). Therefore, particularly in the case of high predicted scores, the predicted score seems relatively accurate. We used target sites with a predicted suppressive activity of >0.75 for construction of the siRNA expression library.

Confirmation of specificity of target sequence

When we select a certain target site in a specific gene, we have to check the specificity (uniqueness) of the sequence of the target site. In other words, we must confirm that there are no highly homologous sequences in other genes. It is necessary to choose the conditions under which such specificity is examined. If very stringent conditions are applied, for example, if more than three base mismatches are needed for discrimination between the correct target and similar targets, a significant fraction of favorable target sites that can be predicted to be highly vulnerable targets can be eliminated using the BLAST search program.

In contrast, under less stringent conditions, for example, if only one base mismatch is allowed for discrimination, the designed siRNA might have potential to disrupt not only the correct target gene but also other homologous genes, even though cleavage efficiency would be reduced in the latter case. We decided to use relatively low-stringency conditions for our examination of target site specificity, as many putative candidate genes can be analyzed in further detail by creating new siRNAs targeted to additional sites within the candidate gene.

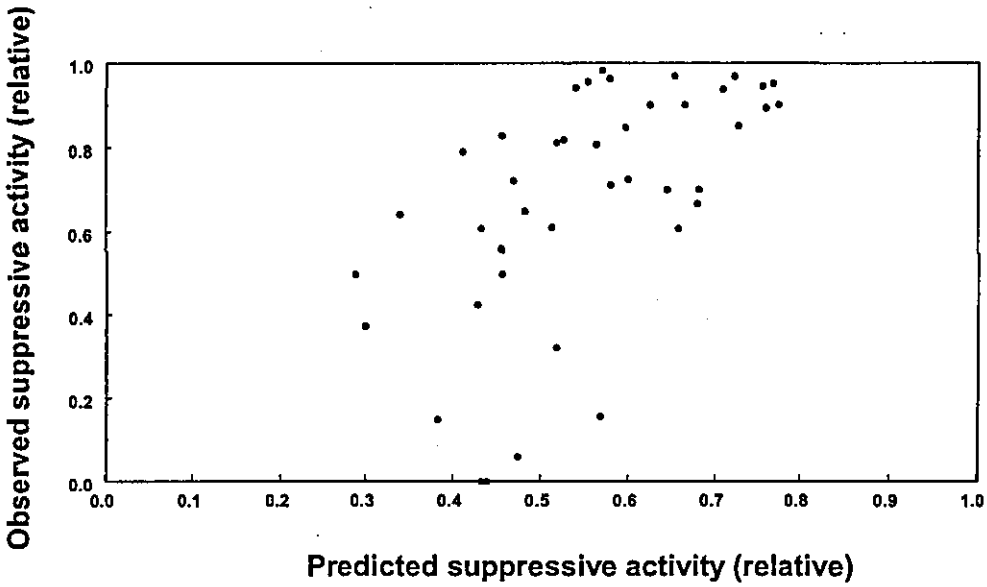

FIG. 3. Relationship between predicted and observed suppressive activities. The actual values were measured in cotransfection experiments with HeLa cells and 10 ng pRL-RSV, 30 ng pGL3, and 10 nM individual siRNA oligonucleotides targeted to the firefly gene for luciferase. The activities of firefly luciferase were normalized by reference to those of *Renilla* luciferase, and suppressive activity was calculated relative to the value obtained with the negative control. Predicted activities were calculated from the target sequence of the siRNA with a nonlinear algorithm, as described in Materials and Methods.

Alternative splicing and highly homologous families of genes also create major problems in selection of the target site. Recent genomewide analysis of transcripts has revealed that many genes yield alternatively spliced transcripts. Thus, for example, in cases where the goal is to disrupt a specific transcript, an siRNA expression vector targeted against that specific transcript can be created without potential damage to other alternatively spliced transcripts that originate from the same prespliced transcript. For generation of an initial library, however, with the goal of targeting a gene that yields several differently spliced mRNAs, the sequence common to all spliced versions of the original transcript (or the sequence common to all members of a family of homologous genes) can be chosen as the target site.

This strategy provides an economical scheme for creation of the first library, and additional siRNAs can be generated easily if the targeted gene turns out to be an attractive candidate.

Improved method for construction of a library

Efficiency, speed, and cost should not be ignored in efforts to construct an siRNA expression library directed against all genes in a genome. Bearing these issues in mind, we have established a large-scale method for constructing libraries (Fig. 4). Specifically, we synthesized 96 sets of oligonucleotides that included hairpin sequences corresponding to 96 target sequences and annealed each sense/antisense set separately. Then, we mixed the annealed oligonucleotides in one tube and ligated them into the *Bsp*MI site of the U6-driven siRNA expression vector. After transformation of *E. coli*, we picked 384 clones and sequenced them to identify each clone.

Of the 384 clones, 358 had the insert sequences that we expected. Therefore, the percentage of cloning failure was 6.8%. The representation of the 96 inserts were as follows: 5 missed, 3 mutated inserts (1 clone for 2 inserts, 2 clones for 1 insert), 1 clone for 9 inserts, 2 clones for 19 inserts, 3 clones for 18 inserts, 4 clones for 11 inserts, 5 clones for 12 inserts, 6 clones for 8 inserts, 7 clones for 3 inserts, >7 clones for 8 inserts. As a result, almost 90% of all anticipated clones were recovered in this single procedure. The sequences corresponding to unrecovered clones were used in a second round of construction. This bulk procedure allowed us to make siRNA expression vectors much more cheaply and rapidly than the traditional procedure, in which each clone is made separately. Using this procedure, we are able to generate 3000–5000 siRNA expression vectors per month (Taira and Miya-gishi, 2001).

Comparison between siRNA expression libraries and siRNA oligonucleotide libraries

Plasmid-based siRNA expression libraries and oligonucleotide-based siRNA libraries each have advantages

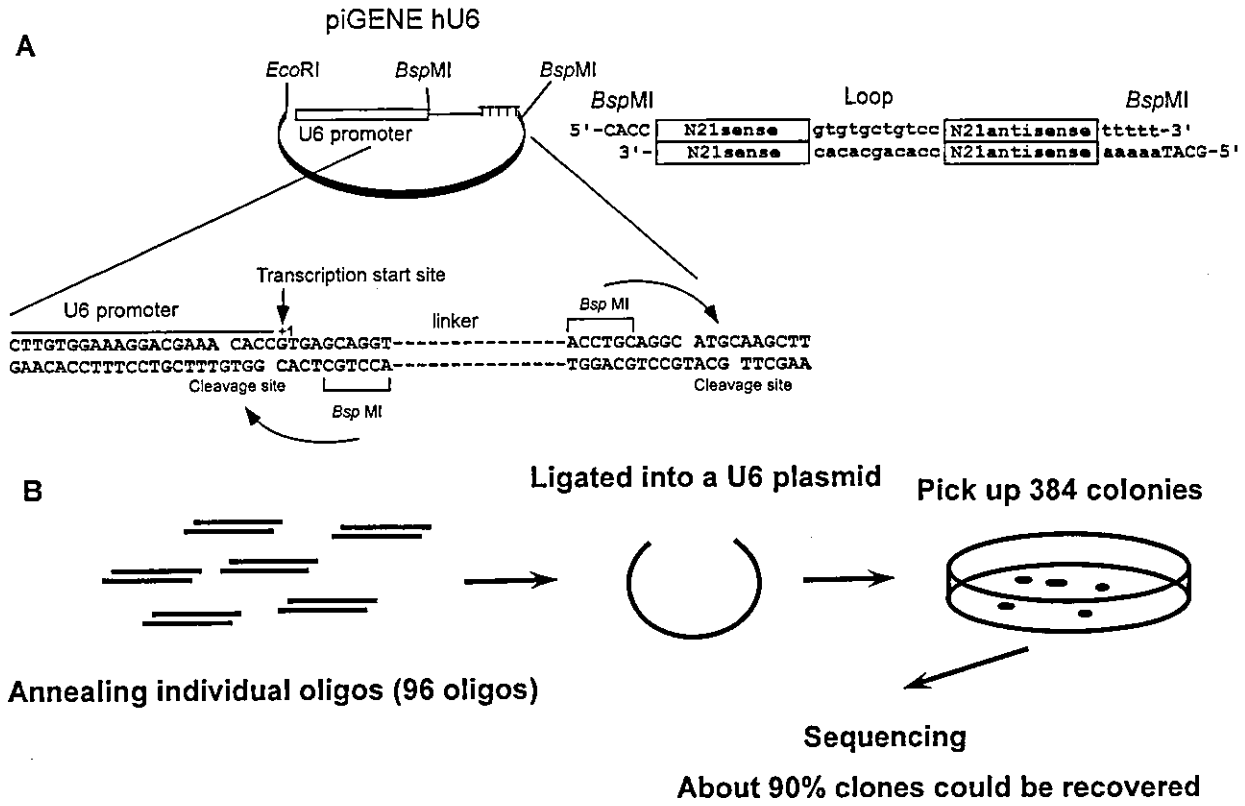

FIG. 4. (A) Cloning site and the sequence of the insert in the hairpin type siRNA expression vector. (B) Strategy for construction of our siRNA expression library.

and disadvantages. The efficiency of transfection of cells with siRNA depends on the type of cell, and RNAi seems to be sustained for only a limited period of time. The advantage of plasmid-based siRNA is that it is possible to eliminate those cells that have not been transfected with plasmids by selection for antibiotic-resistance genes. Moreover, RNAi continues for much longer periods when plasmid-based siRNAs are used. Viral vectors allow delivery of siRNA expression cassettes into cells at high efficiencies of transfection, and in the case of lentivirus and retrovirus, it is easy to generate stable knockdown cells via integration of the viral vector into the genome. It is possible that when siRNA expression vectors are used as a bulk library, viral vectors that can generate stable knockdown cells will prove to be the most suitable for screening with the bulk library.

The advantage of oligonucleotide-based siRNA is that in some lines of cells, the transfection efficiency is >90% and is higher than that obtained with plasmids. However, once plasmid-based siRNAs have been generated, they can be amplified to an unlimited extent, particularly when the construct shown in Figure 2B is used, that is, a construct that does not undergo unwanted mutation (unlike conventional constructs) during amplification in *E. coli*. Such a construct permits the lower-cost generation library.

In conclusion, in the present study, we developed a genetically stable and very active siRNA expression vector and established a system for efficient construction of a high-quality siRNA expression library with potentially genomewide application (the libraries become available through iGENE Therapeutics, Inc., *www.iGENE-therapeutics.co.jp*). We now need to develop a screening system and to identify the functional gene(s) that should be our first targets. Within the next few years, it should be possible to identify a number of genes using an siRNA library. We have been able to identify miRNA sequences using our ribozyme library (unpublished observations). Because miRNA plays an important role in development and perhaps in other biologic phenomena in mammalian cells (Kawasaki and Taira, 2003), siRNA libraries should also include miRNA and undefined noncoding RNAs as targets. Such libraries are likely to shed light on many important normal and abnormal biologic processes.

REFERENCES

BEGER, C., PIERCE, L.N., KRUGER, M., MARCUSSON, E.G., ROBBINS, J.M., WELCSH, P., WELCH, P.J., WELTE, K., KING, M.C., BARBER, J.R., and WONG-STAAL, F. (2001). Identification of Id4 as a regulator of BRCA1 expression by using a ribozyme library-based inverse genomics approach. Proc. Natl. Acad. Sci. USA **98,** 130–135.

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.

CAPLEN, N.J., PARRISH, S., IMANI, 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.

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.

FIRE, A., XU, S., MONTGOMERY, M.K., KOSTAS, S.A., DRIVER, S.E., and MELLO, C.C. (1998). Potent and specific gene interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391,** 806–811.

FRASER, A.G., KAMATH, R.S., ZIPPERLEN, P. MARTINEZ-CAMPOS, M., SOHRMANN, M., and AHRINGER, J. (2000). Functional genomic analysis of *C. elegans* chromosome 1 by systemic RNA interference. Nature **408,** 225–230.

GONCZY, P., ECHEVERN, C. OEGEMA, K. COULSON, A., JONES, S.J., COPLEY, R.R., DUPERON, J., OEGEMA, J., BREHM, M., CASSIN, E., HANNAK, E., KIRKHAM, M., PICHLER, B., FLOHRS, K., GOESSEN, A., LEIDEL, S., ALLEAUME, A.M., MARTIN, C., OZLU, N., BORK, P., AND HYMAN, A.A. (2000). Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature **408,** 331–336.

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.

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.

KAWASAKI, H., and TAIRA, K. (2003). Short hairpin type of dsRNAs that are controlled byt RNA^{Val} promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. Nucleic Acids Res. **31,** 700–707.

KETTING, R.F., FISCHER, S.E., BERNSTEIN, E., SIJEN, T., HANNON, G.J., and PLASTERK, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *Caenorhabditis elegans*. Genes Dev. **15,** 2654–2659.

KRUGER, M., BERGER, C., LI, Q.X., WELCH, P.J., TRITZ, R., LEAVITT, M., BARBER, J.R., and WONG-STAAL, F. (2000). Identification of eIF2Bgamma and eIF2gamma as cofactors of hepatitis C virus internal ribosome entry site-mediated translation using a functional genomics approach. Proc. Natl. Acad. Sci. USA **97,** 8566–8571.

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.

LI, Q.X., ROBBINS, J.M., WELCH, P.J., WONG-STAAL, F., and BARBER, J.R. (2000). A novel functional genomics approach identifies mTERT as a suppressor of fibroblast transformation. Nucleic Acids Res. **28,** 2605–2612.

LINDBERG, W., PERSSON, J.A., and WOLD, S. (1983). Partial least-squares method for spectrofluorimetric analysis of mixtures of humic acid and lignin sulfonate. Anal. Chem. **55,** 643–648.

McMANUS, M.T., and SHARP, P.A. (2002). Gene silencing in mammals by small interfering RNAs. Nat. Rev. Genet. **3,** 737–747.

MIYAGISHI, M., FUJII, R., HATTA, M., YOSHIDA, E., ARAYA, N., NAGAFUCHI, A., ISHIHARA, S., NAKAJIMA, T., and FUKAMIZU, A. (2001). Regulation of Lef-mediated transcription and p53-dependent pathway by associating β-catenin with CBP/p300. J. Biol. Chem. **275,** 35170–35175.

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.

MIYAGISHI, M., and TAIRA, K. (2003). Expression of siRNA from a pol III promoter in mammalian cells. In: *Perspectives in Gene Expression.* K. Appasani, ed. (The Eaton Publishers, Westboro, MA), pp. 361–376.

ONUKI, R., NAGASAKI, A., KAWASAKI, H., BABA, T., UYEDA, T.Q., and TAIRA, K. (2002). Confirmation by FRET in individual living cells of the absence of significant amyloid beta-mediated caspase 8 activation. Proc. Natl. Acad. Sci. USA **99,** 14716–14721.

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.

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.

RHOADES, K., and WONG-STAAL, F. (2003). Inverse genomics as a powerful tool to identify novel targets for the treatment of neurodegenerative diseases. Mech. Ageing Dev. **124,** 125–132.

SUI, G., SOOHOO, C., AFFAREL, B., GAY, F., SHI, Y., FORRESTER, W.C., and SHI, Y. (2002). A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proc. Natl. Acad. Sci. USA **99,** 5515–5520.

SUYAMA, E., KAWASAKI, H., KASAOKA, T., and TAIRA, K. (2003a). Identification of genes responsible for cell migration by a library of randomized ribozymes. Cancer Res. **63,** 119–124.

SUYAMA, E., KAWASAKI, H., NAKAJIMA, M., and TAIRA, K. (2003b). Identification of genes involved in cell invasion by using a library of randomized hybrid ribozymes. Proc. Natl. Acad. Sci. USA **100,** 5616–5621.

TAIRA, K., and MIYAGISHI, M. (2001). siRNA expression system and method for producing functional gene knock-

down cell using the system. Japanese Patent Application, 2001-363385.

TUSCHL, T. (2002). Expanding small RNA interference. Nat. Biotechnol. **20,** 446–448.

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.

WELCH, P.J., MARCUSSON, E.G., LI, Q.X., BEGER, C., KRUGER, M., ZHOU, C., LEAVITT, M., WONG-STAAL, F., and BARBER, J.R. (2000). Identification and validation of a gene involved in anchorage-independent cell growth control using a library of randomized hairpin ribozymes. Genomics **66,** 74–83.

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.

ZAMORE, P.D. (2001). RNA interference: Listening to the sound of silence. Nat. Struct. Biol. **8,** 746–750.

Address reprint requests to:
Prof. Kazunari Taira
Department of Chemistry and Biotechnology
School of Engineering
University of Tokyo
Hongo, Tokyo 113-8656
Japan

E-mail: makoto-m@sannet.ne.jp

Received July 5, 2003; accepted in revised form August 8, 2003.

increased frequency of meiotic and/or inherited late somatic recombination.

It will be interesting to elucidate further the relationship of SRS with SAR signalling as well as with systemic wound signalling[19], systemic post-transcriptional gene silencing[20] and systemic acquired acclimatization to light[21]. The genomic changes reported here might constitute an adaptive measure in response to biotic stress. Of special utility may be recombination events that lead to new specificities in pathogen resistance genes[22]. □

Methods

Generation of transgenic plants

The generation and analysis of luciferase recombination test plants has been described[23].

Plant inoculation

Single leaves of 6–8-week-old TMV-sensitive Havana 425 tobacco plants were rub-inoculated with 300 ng of full-length infectious TMV or ORMV RNA transcripts as described[16]; control plants were mock-inoculated with phosphate buffer. After treatment, plants were kept at 22 °C or 32 °C and recombination frequencies were scored 7 d after inoculation. In another experiment, single TMV-inoculated leaves of plants inoculated at 32 °C were excised and removed from the plants at times between 0 and 36 h (5 min, 8 h, 24 h, 36 h) or left on the plants. We detected disease symptoms on non-inoculated leaves 10 d after inoculation. Increase of recombination (ratio to control) was analysed in non-inoculated leaves 7 d after inoculation.

Assay for infection

The presence or absence of the viral RNA was tested by polymerase chain reaction with reverse transcription (RT–PCR; primers: forward, 5′-CTGGTGAAGTATTTGTCTGA-3′, and reverse, 5′-ACCCGCTGACATCTTCACAT-3′) done on the remaining tissue 2 weeks after inoculation (Fig. 1c).

Grafting experiments

We inoculated single leaves of several 10-week-old Havana 425 tobacco plants with TMV at 22 °C or 32 °C, and after 24 h we grafted upper, virus-free leaves from these plants onto 10-week-old healthy plants. The scion leaves were grafted onto the sides of the stem in place of the removed third leaf (Fig. 2a). After grafting, the plants were kept at either 22 °C or 32 °C. We measured the recombination frequency in the newly emerged leaves 10–14 d after grafting.

Su/su tobacco plants were infected with 300 ng of ORMV RNA. Grafting experiments similar to those done with TMV were conducted. We scored dark green and albino sectors on newly emerged tissue 1 month after grafting (Fig. 2d).

Statistical treatment of the data

Averages and standard deviations were calculated in all experiments. The statistical significance of the experiments was confirmed by Student's *t*-test (two-tailed paired or non-paired) and the Fisher test.

Received 20 December 2002; accepted 14 April 2003; doi:10.1038/nature01683.

1. Puchta, H., Swoboda, P. & Hohn, B. Induction of homologous DNA recombination in whole plants. *Plant J.* **7**, 203–210 (1995).
2. Kovalchuk, I., Kovalchuk, O., Arkhipov, A. & Hohn, B. Transgenic plants are sensitive bioindicators of nuclear pollution caused by the Chernobyl accident. *Nature Biotechnol.* **16**, 1054–1057 (1998).
3. Ries, G. *et al.* Elevated UV-B radiation reduces genome stability in plants. *Nature* **406**, 98–101 (2000).
4. Ries, G., Buchholz, G., Frohnmeyer, H. & Hohn, B. UV-damage-mediated induction of homologous recombination in *Arabidopsis* is dependent on photosynthetically active radiation. *Proc. Natl Acad. Sci. USA* **97**, 13425–13429 (2000).
5. Walbot, V. Reactivation of Mutator transposable elements of maize by ultraviolet light. *Mol. Gen. Genet.* **234**, 353–360 (1992).
6. Walbot, V. UV-B damage amplified by transposons in maize. *Nature* **397**, 398–399 (1999).
7. Kumar, A. & Bennetzen, J. L. Plant retrotransposons. *Annu. Rev. Genet.* **33**, 479–532 (1999).
8. Dellaporta, S. *et al.* Endogenous transposable elements associated with virus infection in maize. *Cold Spring Harbor Symp. Quant. Biol.* **49**, 321–328 (1984).
9. Johns, M. A., Mottinger, J. & Freeling, M. A low copy number, copia-like transposon in maize. *EMBO J.* **4**, 1093–1102 (1985).
10. Brakke, M. Mutations, the aberrant ratio phenomenon, and virus infection of maize. *Annu. Rev. Phytopathol.* **22**, 77–94 (1984).
11. Lucht, J. M. *et al.* Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nature Genet.* **30**, 311–314 (2002).
12. Harrison, B. D. & Wilson, T. M. A. Milestones in the research on tobacco mosaic virus. *Phil. Trans. R. Soc. Lond. B* **354**, 521–529 (1999).
13. Aguilar, I., Sanchez, F., Martin, A., Martinez-Herrera, D. & Ponz, F. Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic virus), a crucifer tobamovirus infecting *Arabidopsis thaliana*. *Plant Mol. Biol.* **30**, 191–197 (1996).
14. Burk, L. G. & Menser, H. A. A dominant *aurea* mutation in tobacco. *Tob. Sci.* **8**, 101–104 (1964).
15. Shalev, G., Sitrit, Y., Avivi-Ragolski, N., Lichtenstein, C. & Levy, A. A. Stimulation of homologous recombination in plants by expression of the bacterial resolvase RuvC. *Proc. Natl Acad. Sci. USA* **96**, 7398–7402 (1999).
16. Boyko, V., Ferralli, J., Ashby, J., Schellenbaum, P. & Heinlein, M. Function of microtubules in intercellular transport of plant virus RNA. *Nature Cell Biol.* **2**, 826–832 (2000).
17. Dempsey, D. A., Shah, J. & Klessig, D. Salicylic acid and disease resistance in plants. *Crit. Rev. Plant Sci.* **18**, 547–575 (1999).
18. Dong, X. Genetic dissection of systemic acquired resistance. *Curr. Opin. Plant Biol.* **4**, 309–314 (2001).
19. Pearce, G., Strydom, D., Johnson, S. & Ryan, C. A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. *Science* **253**, 895–897 (1991).
20. Waterhouse, P., Wang, M.-B. & Finnegan, J. Role of short RNAs in gene silencing. *Trends Plant Sci.* **6**, 297–301 (2001).
21. Karpinski, S. *et al.* Systemic signalling and acclimation in response to excess excitation energy in *Arabidopsis*. *Science* **284**, 654–657 (1999).
22. Whitham, S. *et al.* The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. *Cell* **78**, 1101–1115 (1994).
23. Gorbunova, V. *et al.* A new hyperrecombinogenic mutant of *Nicotiana tabacum*. *Plant J.* **24**, 601–611 (2000).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank S. van Eeden for technical assistance and H. Rothnie, K. Smith and E. Schultz for comments on the manuscript. The Novartis Research Foundation and Alberta Ingenuity Grant are acknowledged for financial support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to I.K. (igor.kovalchuk@uleth.ca).

Cloning of adiponectin receptors that mediate antidiabetic metabolic effects

Toshimasa Yamauchi*†‡, Junji Kamon*‡§, Yusuke Ito*, Atsushi Tsuchida*, Takehiko Yokomizo||, Shunbun Kita*, Takuya Sugiyama¶, Makoto Miyagishi#, Kazuo Hara*†, Masaki Tsunoda☆, Koji Murakami☆, Toshiaki Ohteki†, Shoko Uchida*, Sato Takekawa*, Hironori Waki*†, Nelson H. Tsuno††, Yoichi Shibata††, Yasuo Terauchi*†, Philippe Froguel‡‡, Kazuyuki Tobe*†, Shigeo Koyasu**†, Kazunari Taira#, Toshio Kitamura¶, Takao Shimizu||, Ryozo Nagai* & Takashi Kadowaki*†**

** Department of Internal Medicine, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan*
† CREST of Japan Science and Technology Corporation, 332-0012, Japan
§ Biological Research Laboratories, Nissan Chemical Industries, Saitama 349-0294, Japan
|| Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Tokyo, Tokyo 113-0033, CREST and PRESTO of JST, Japan
¶ Division of Hematopoietic Factors, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan
 # Department of Chemistry and Biotechnology, School of Engineering, University of Tokyo, Tokyo 113-8656, and Gene Function Research Center, National Institute of AIST, Tsukuba 305-8562, Japan
☆ Central Research Laboratories, Kyorin Pharmaceutical, Tochigi 329-0114, Japan
*** Department of Microbiology and Immunology, Keio University School of Medicine, Tokyo 160-8582, Japan*
†† Department of Transfusion Medicine, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan
‡‡ Institute of Biology-CNRS, Pasteur Institute of Lille, UPRES A8090, 59000 Lille, France
‡ These authors contributed equally to this work

Adiponectin (also known as 30-kDa adipocyte complement-related protein; Acrp30)[1–4] is a hormone secreted by adipocytes that acts as an antidiabetic[5–12] and anti-atherogenic[8,12,13] adipokine. Levels of adiponectin in the blood are decreased under conditions of obesity, insulin resistance and type 2 diabetes[2]. Administration of adiponectin causes glucose-lowering effects