using a DC Protein Assay Kit (Bio-Rad) with bovine serum albumin as the standard protein.

For western blotting analysis, samples (10 μg total protein) were fractionated by SDS–PAGE (15% polyacrylamide) and bands of proteins were transferred to an Immobilon™-P^{SQ} transfer membrane (Millipore, Billerica, CA). After blocking in TBS-T (Tris-buffered saline containing 0.1% Tween) containing 3% skimmed milk, the membrane was incubated for 1 h with polyclonal antibodies raised in rabbit against Cre recombinase (1:1000 dilution in TBS-T plus 0.3% skimmed milk) (Novagen) and for 1 h with antibodies against rabbit Ig conjugated to horseradish peroxidase (1:1000 dilution in TBS-T plus 0.3% skimmed milk) (Amersham Biosciences, Little Chalfont, UK). As a control for the separation of nuclear and cytosolic fractions, we also performed western blotting by incubating for 1 h with monoclonal antibodies raised in mouse against β-tubulin (anti-β-tubulin, clone AA2, 1:1000 dilution in TBS-T plus 0.3% skimmed milk) (Upstate) and 1 h with antibodies against mouse Ig conjugated to horseradish peroxidase (1:1000 dilution in TBS-T plus 0.3% skimmed milk) (Amersham Biosciences). The results of western blotting are shown in Figure 2C.

Amplification by PCR of transfected plasmids

Cells were grown in 24-well tissue culture plates, transfected with 375 ng iGL3B Cre-On plasmid, 50 ng pGL3 and 15 ng pRL-RSV and incubated with TAT–NLS–Cre (30 μg/ml) or dialysis medium as described above. Then, 22 h later, cells were resuspended in Passive Lysis Buffer (Promega) and centrifuged (3000 r.p.m., 10 min, 4°C). After two washes with Tris–EDTA buffer (TE) [10 mM Tris–HCl (pH 8.0) plus 1 mM EDTA], the pellets were resuspended in TE and boiled (65°C for 10 min and 99°C for 5 min). Then 4–6 μl of each sample was used as template for PCR, which was performed with Ex Taq™ (Takara Shuzo, Shiga, Japan). The results are shown in Figure 3.

RESULTS

In this study, our goal was to establish an exogenously Cre controllable U6 promoter-based siRNA expression vector. Since Cre recombinase catalyzes the excision of sequences between *loxP* sites, our strategy was to place, between two *loxP* sites, an insert that inhibited expression of complete siRNAs to generate a Cre-On system. Schematic representatives of the switches for the expression of siRNAs are shown in Figure 1. In this scheme, two *loxP* sites (with a fragment of 816 bp inserted between them) are placed between sense and antisense regions of the siRNA encoding sequence. The 816 bp insert starts with TTTTTTT, which causes termination of transcription, since TTTT is a terminator sequence for the U6 promoter in mammalian cells. As a result, only sense RNAs, which have no suppressive effect, can be transcribed in the absence of Cre recombinase. In the presence of Cre recombinase, recombination occurs and the inserted fragment, including the termination sequence, is excised. The excision of this fragment allows transcription of full, functional siRNAs since, now, only a 34 bp *loxP* site connects the sense and antisense sequences, yielding short hairpin RNAs (shRNAs) (Fig. 1).

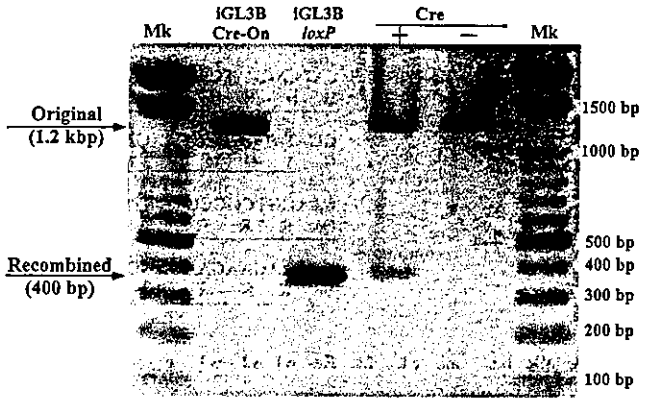

Figure 3. Fractionation by gel electrophoresis of products of PCR amplified from iGL3B*loxP* (positive control), iGL3BCre-On (negative control), Cre + (nuclei from cells transfected with iGL3BCre-On and incubated with TAT-NLS-Cre, final concentration 30 μg/ml) and Cre – (nuclei from cells transfected with iGL3BCre-On and incubated with dialysis medium). Mk, size markers. Products of PCR from cells incubated with TAT-NLS-Cre protein yielded both 'Original' and 'Recombined' bands, while cells incubated with dialysis medium yielded only the 'Original' band. These results indicate that recombination occurred only in the presence of the exogenously added TAT-NLS-Cre.

Construction of Cre recombinase controllable siRNA expression vectors

To study the possibility of controlling the expression of siRNA with Cre recombinase, we used three types of siRNA expression vector directed against firefly luciferase, namely iGL3B, iGL3B*loxP* and iGL3BCre-On, the first two of which were used as positive controls. An siRNA expression vector directed against DsRed2 (pU6iRed) was also used as an unrelated negative control. iGL3B is a vector that produces 21 nt hairpin-type siRNA with a 9 nt loop, constructed in our laboratory as described previously (41). We designed iGL3BCre-On and iGL3B*loxP* on the basis of the sense and antisense sequences of iGL3B. iGL3BCre-On was designed for the Cre-On system (Fig. 1) and iGL3B*loxP* was similar to iGL3B, with the exception that the 9 nt loop was replaced by a 34 nt *loxP* sequence. In other words, iGL3B*loxP* had exactly the same structure as that of recombined iGL3BCre-On (Fig. 1, after recombination). This vector was constructed to examine the efficiency with which recombined iGL3BCre-On induced RNAi against the targeted gene for firefly luciferase. A comparison of efficiency between iGL3B and iGL3B*loxP* is shown below in Figure 4A (compare lanes 5 and 6). As discussed below, at least at the concentration of plasmid used, the efficiencies of iGL3B and iGL3B*loxP* were almost the same. However, at a lower concentration of plasmids, iGL3B*loxP* with a larger 34 nt loop had a significantly lower activity than iGL3B with a 9 nt loop, indicating that the size of the loop within the shRNAs has some influence on the activity.

We chose the firefly gene for luciferase as the target of siRNA for the following reasons. First, the efficiency of gene silencing by siRNA depends to a large extent on the choice of target site. In this study, our siRNA expression vectors were directed against the target site of iGL3B, which has been shown to be very susceptible to RNAi. Since we knew that we would be unlikely to achieve 100% recombination, we hoped

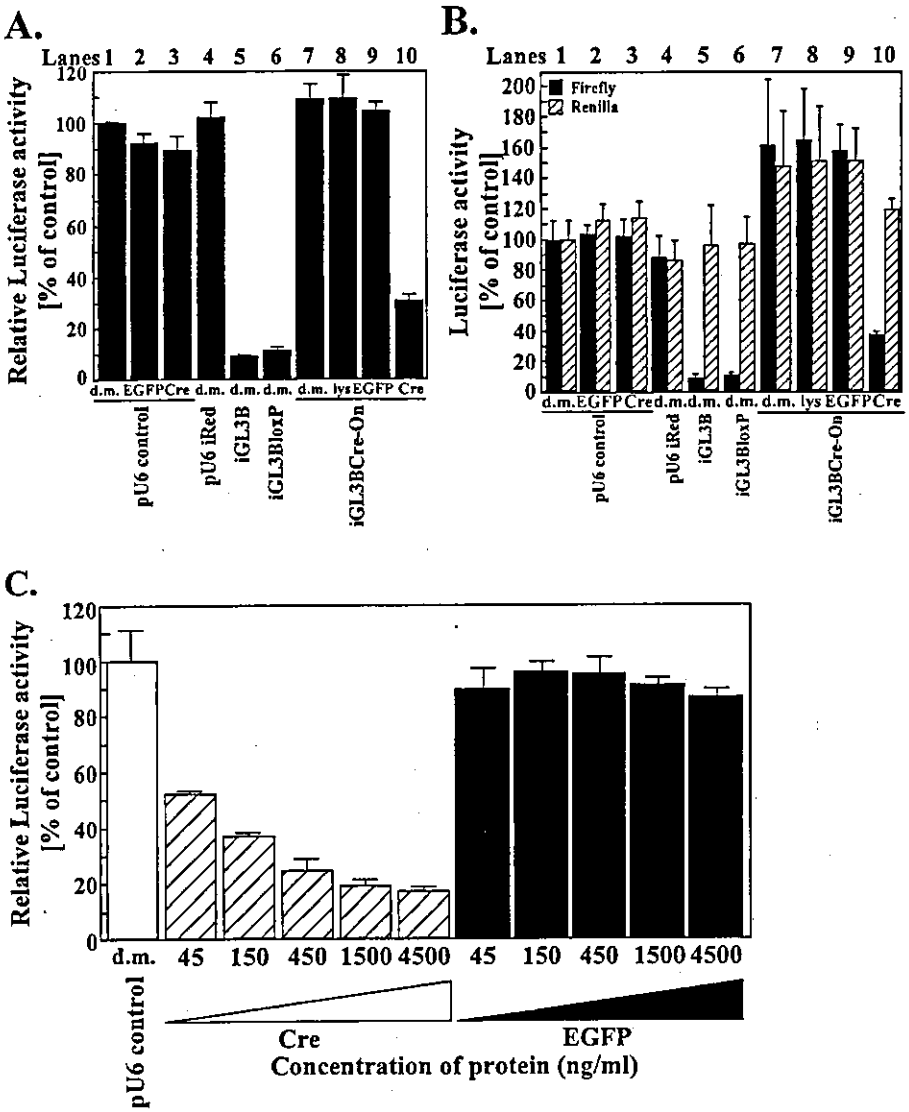

Figure 4. (A) Luciferase activity due to iGL3BCre-On. Lane 1, negative control (d.m., dialysis medium); lane 2, pU6 control and 4.5 μg/ml TAT-NLS-EGFP (EGFP); lane 3, pU6 control and 4.5 μg/ml TAT-NLS-Cre (Cre); lane 4, pU6iRed and dialysis medium; lane 5, iGL3B and dialysis medium; lane 6, iGL3B*loxP* and dialysis medium; lane 7, iGL3BCre-On and dialysis medium; lane 8, iGL3BCre-On and residual *E.coli* lysate (lys); lane 9, iGL3BCre-On and 4.5 μg/ml TAT-NLS-EGFP; lane 10, iGL3BCre-On and 4.5 μg/ml TAT-NLS-Cre. Compared with the result for the negative control (pU6 control and incubation with dialysis medium), the luciferase activity in the cells transfected with iGL3BCre-On and incubated with TAT-NLS-Cre was suppressed by ~70%, while incubation with dialysis medium, with TAT-NLS-EGFP or with residual *E.coli* lysate did not suppress luciferase activity. (B) Firefly and *Renilla* luciferase activities of the results shown in (A). The activity of *Renilla* was not affected by the addition of recombinant protein, dialysis medium or residual *E.coli* lysate. (C) The extent of suppression of luciferase activity decreased with decreases in the concentration of TAT-NLS-Cre, while incubation with TAT-NLS-EGFP at various concentrations failed to have any effect on luciferase activity, which was the same as the control (incubated with dialysis medium). The results were calculated by normalizing the activity due to firefly luciferase to that due to *Renilla* luciferase. Then the results were calculated as percentages of the control value. Each bar indicates an average value and vertical bars indicate standard errors of triplicate assays.

to induce successful RNAi by choosing a very susceptible target, even if not all of the siRNA expression vectors underwent recombination. Second, the choice of firefly luciferase as reporter allowed us to perform quantitative assays using the standard dual luciferase system (11).

Expression and purification of TAT–NLS–Cre

To use NLS–Cre in cells, especially in the nucleus, we decided to deliver the NLS–Cre recombinant protein directly rather than delivering its expression vector. However, in this case it is necessary to enhance the entry of NLS–Cre into cells. It has been reported that NLS can, itself, function as a protein transduction domain (PTD) and promote the cellular uptake of recombinant protein, but the combination of both TAT and NLS is more efficient (34). Therefore, we constructed pTriEx-3 Hygro-TAT-NLS-Cre, which encodes the 41 kDa fusion protein TAT–NLS–Cre recombinase (TAT–NLS–Cre) that includes an N-terminal histidine tag and is expressed in *E.coli* (Fig. 2A). We purified the fusion protein from a lysate of *E.coli* cells by Ni chelate affinity chromatography and checked the purity of the protein by SDS–PAGE (Fig. 2B, 41 kDa).

Confirmation of the localization and activity of TAT–NLS–Cre

To determine whether the TAT–NLS–Cre protein could be used to control the expression of siRNAs by the strategy described above, we first confirmed the localization and ability of TAT–NLS–Cre to catalyze recombination. After incubating HeLa S3 cells in medium supplemented with the TAT–NLS–Cre protein for the indicated times, we washed the cells twice with PBS, incubated them with trypsin in order to eliminate TAT–NLS–Cre attached to cell membranes and prepared cytosolic and nuclear fractions. Then we performed western blotting of these fractions with antibodies against Cre recombinase to confirm the localization of TAT–NLS–Cre. The results are shown in Figure 2C. It appeared that a 1 h incubation was sufficient to allow entry of TAT–NLS–Cre into nuclei. Incubation for 1 h increased the concentration of TAT–NLS–Cre in nuclei but no further increase was apparent after incubation for 2 h. TAT–NLS–Cre was detected in the nuclear fraction, whereas β-tubulin, which was used as a control for separation of the nuclear and cytosolic fractions, was detected in the cytosolic fraction, indicating that there was no detectable cytosolic contamination in the nuclear fraction.

To examine whether TAT–NLS–Cre retained the activity of Cre recombinase and could catalyze the recombination of *loxP* sites, we transfected HeLa S3 cells with iGL3BCre-On and incubated them with TAT–NLS–Cre or with dialysis medium under the conditions indicated above. Plasmids were then extracted from the cells and amplified by PCR. If recombination had occurred, the extracted plasmids should have had exactly the same structure as that of iGL3B*loxP*, yielding a PCR product of ~400 bp (indicated by 'Recombined' in Fig. 3). If recombination had not occurred, the extracted plasmids should have retained the initial structure, yielding a PCR product of ~1.2 kb (indicated by 'Original' in Fig. 3). As shown in Figure 3, PCR analysis of cells incubated with dialysis medium yielded a fragment of 1.2 kb, while PCR analysis of cells incubated with TAT–NLS–Cre yielded both 1.2 kb and 400 bp fragments. From these results we conclude that although we had not achieved a recombination efficiency of 100%, TAT–NLS–Cre did have the ability to enter cells, to find its way to the nucleus and to catalyze recombination.

The suppressive effect of iGL3BCre-On can be switched on by NLS–Cre in a dose-dependent manner

Finally, we examined whether we could control the expression of iGL3BCre-On using TAT–NLS–Cre recombinant protein. Four hours after co-transfecting HeLa S3 cells with siRNA expression vectors or the mock vector (pU6 control) together with expression vectors for firefly luciferase and *Renilla* luciferase, we added TAT–NLS–Cre or, as a control, TAT–NLS–EGFP, residual *E.coli* lysate, obtained by the same procedure as TAT–NLS–Cre, or dialysis medium (for the dialysis of TAT–NLS–Cre) to the medium. We then analyzed luciferase activities on the following day and calculated the luciferase activities relative to those obtained with the pU6 control (Fig. 4A, lane 1). As shown in Figure 4A, incubation of cells transfected with iGL3BCre-On with TAT–NLS–Cre (lane 10) resulted in ~70% suppression of firefly luciferase activity.The corresponding suppression obtained with positive controls was 80–90% (Fig. 4A, lanes 5 and 6).

In contrast, the corresponding set of incubation experiments indicated that the addition of TAT–NLS–Cre to the pU6 empty vector (negative control) did not cause any suppression (lanes 1–3). Similarly, the addition of TAT–NLS–EGFP (lane 9) or residual *E.coli* lysate (lane 8) to the effector plasmid, iGL3BCre-On, showed no effect. Moreover, cells transfected with pU6iRed (an unrelated siRNA expression vector against DsRed2, lane 4) also showed no suppressive effect, indicating that the suppression of expression of the target gene was not due to a non-specific effect of dsRNA or TAT–NLS–Cre itself.

Finally, we examined whether the dose of TAT–NLS–Cre affects the switching of iGL3BCre-On by incubating iGL3BCre-On-transfected HeLa S3 cells with TAT–NLS–Cre at various concentrations. As shown in Figure 4C, increasing the concentration of TAT–NLS–Cre increased the efficiency of switching of iGL3BCre-On. We also incubated iGL3BCre-On-transfected HeLa S3 cells with TAT–NLS–EGFP at various concentrations as negative controls. TAT–NLS–EGFP had no suppressive effects on luciferase activity at any of the tested concentrations.

Taken together with the fact that TAT–NLS–Cre could successfully enter the nucleus and cause recombination between *loxP* sites, the present results clearly demonstrate that iGL3BCre-On had been switched on by the exogenously supplied TAT–NLS–Cre recombinant protein.

DISCUSSION

We have here described one of the first examples of the control of the expression of siRNA in a pol III system by Cre recombinase (43) and the first example using TAT–NLS–Cre recombinase fusion protein. We have demonstrated that our Cre-On siRNA expression vector was switched on only in the presence of Cre recombinase and that the gene suppressive activity depended on the concentration of Cre recombinase.

The use of Cre recombinase to control the expression of siRNA has several advantages. First, Cre recombinase can be expressed stably, and both tissue type- and cell type-specifically and its expression is heritable in mammals, such as mice (32,33,44,45). Thus, it should be possible to control the expression of siRNA in a tissue-specific manner using a pol III promoter such as the U6 promoter, whose expression is constitutive and ubiquitous. It has been reported that siRNA can also be expressed using a pol II system (22,46), which might allow tissue-specific expression. However, a pol III system might have some advantages as a consequence of its high level of activity (~4 × 10^5 copies of the pol III transcript per cell) (47). Second, the expression of Cre can be regulated by tetracycline (45), a hormone (48), a steroid (49) and an interferon (50), if the Cre coding region is fused to a ligand-binding domain specific for the corresponding compound. Such regulation of expression of Cre would allow control of the level of expression of Cre and, as a consequence, of siRNA.

Transgenic Cre recombinase can be expressed in cells and organisms, but problems do occur. In this study, we first tried to express NLS–Cre encoded by pEF9-NLS-Cre inside HeLa S3 cells and to control the expression of the Cre-On siRNA expression vector, but we failed to observe the expected results. Possible reasons for this failure are as follows.

Although the results presented in this work are highly reproducible, the results of co-transfection experiments using reporter plasmids and a NLS–Cre-expressing plasmid gave much larger error bars (not shown). Thus, *trans* effects between the pol II promoters of NLS–Cre (the EF-1α promoter) and reporters (pGL3 and pRL-RSV) on co-transfected plasmids had probably occurred (51).

Furthermore, since time is required for the expression of NLS–Cre and recombination of siRNA plasmids, it is possible that the reporter proteins (firefly and *Renilla* luciferases) might have been expressed before recombination could take place, masking any effects due to switching on the expression of siRNA. The reporter proteins might remain stable for some time and luciferase activities might not have been affected directly, even though no new corresponding mRNAs were produced. It is possible that selection of a suitable promoter for each experimental system and establishment of a cell line or transgenic animal that stably expresses NLS–Cre so that recombination can occur directly after transfection of siRNA expression vectors might overcome these problems. An alternative is to use NLS–Cre fused to the TAT peptide, as we demonstrated in the present study. TAT-fused recombinant proteins move rapidly into a wide range of cells, solving the problem of promoters and the time required for expression of NLS–Cre. Moreover, it is then no longer necessary to establish a suitable stable NLS–Cre-expressing cell line for each experimental system.

In conclusion, we have demonstrated that expression of siRNA under the control of a pol III promoter can be regulated by Cre recombinase. We were able to demonstrate delivery of Cre recombinase into nuclei after incubating cells in medium supplemented with TAT–NLS–Cre. Moreover, delivery should also be possible after establishment of stable Cre recombinase-expressing cell lines or transgenic organisms. The choice of delivery system will depend on the needs of each experiment. Although siRNA technology will not replace the conventional recombination-mediated knockout, it complements the conventional technology because analysis of housekeeping genes might become possible by spatio-temporal control of gene expression, as described in this paper. Collectively, our work has opened up new possibilities for using and controlling the use of siRNAs in research and therapeutics.

ACKNOWLEDGEMENTS

This study was partly supported by grants from the National Institute of Advanced Industrial Science and Technology (AIST), Japan. The authors wish to thank Dr D. Ow, Plant Gene Expression Center, California, for kindly providing plasmid pMM23.

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. 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.
3. Provost,P., Dishart,D., Doucet,J., Frendewey,D., Samuelsson,B. and Radmark,O. (2002) Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J.*, **21**, 5864–5874.
4. 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.
5. Fire,A. (1999) RNA-triggered gene silencing. *Trends Genet.*, **15**, 358–363.
6. Sharp,P.A. (2001) RNA interference 2001. *Genes Dev.*, **15**, 485–490.
7. Hammond,S.M., Caudy,A.A. and Hannon,G.J. (2001) Post-transcriptional gene silencing by double-stranded RNA. *Nat. Rev. Genet.*, **2**, 110–119.
8. Zamore,P.D. (2001) RNA interference: listening to the sound of silence. *Nature Struct. Biol.*, **8**, 746–750.
9. 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.
10. 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.
11. 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.
12. 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.
13. Kawasaki,H. and Taira,K. (2003) Short hairpin type of dsRNAs that are controlled by $tRNA^{Val}$ promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res.*, **31**, 700–707.
14. 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.
15. 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.
16. Sui,G., Soohoo,C., Affar,el 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.
17. 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.
18. 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.
19. Carmell,M.A., Zhang,L., Conklin,D.S., Hannon,J.G. and Rosenquist,T.A. (2003) Germline transmission of RNAi in mice. *Nature Struct. Biol.*, **10**, 91–92.
20. Hasuwa,H., Kaseda,K., Einarsdottir,T. and Okabe,M. (2002) Small interfering RNA and gene silencing in transgenic mice and rats. *FEBS Lett.*, **532**, 227–230.
21. Rubinson,D.A., Dillon,C.P., Kwiatkowski,A.V., Sievers,C., Yang,L., Kopinja,J., Zhang,M., McManus,M.T., Gertler,F.B., Scott,M.L. and Parijs,L.V. (2003) A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nature Genet.*, **33**, 401–406.
22. 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.
23. Miyagishi,M., Sumitomo,H., Miyoshi,H., Kawakami,Y. and Taira,K. (2004) Optimization of an siRNA-expression system with an improved hairpin and its significant suppressive effects in mammalian cells. *J. Gene Med.*, in press.
24. van de Wetering,M., Oving,I., Muncan,V., Pon Fong,M.T., Brantjes,H., van Leenen,D., Holstege,F.C., Brummelkamp,T.R., Agami,R. and Clevers,H. (2003) Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. *EMBO Rep.*, **4**, 609–615.
25. Matsukura,S., Jones,P.A. and Takai,D. (2003) Establishment of conditional vectors for hairpin siRNA knockdowns. *Nucleic Acids Res.*, **31**, e77.

26. Czauderna,F., Santel,A., Hinz,M., Fechtner,M., Durieux,B., Fisch,G., Leenders,F., Arnold,W., Giese,K., Klippel,A. and Kaufmann,J. (2003) Inducible shRNA expression for application in a prostate cancer mouse model. *Nucleic Acids Res.*, **31**, e127.
27. Sternberg,N. and Hamilton,D. (1981) Bacteriophage P1 site-specific recombination. I. Recombination between *loxP* sites. *J. Mol. Biol.*, **150**, 467–486.
28. Abremski,K., Hoess,R. and Sternberg, N (1983) Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell*, **32**, 1301–1311.
29. Sauer,B. (1987) Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **7**, 2087–2096.
30. Odell,J., Caimi,P., Sauer,B. and Russell,S. (1990) Site-directed recombination in the genome of transgenic tobacco. *Mol. Gen. Genet.*, **223**, 369–378.
31. Sauer,B. and Henderson,N. (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl Acad. Sci. USA*, **85**, 5166–5170.
32. Lakso,M., Sauer,B., Mosinger,B.,Jr, Lee,E.J., Manning,R.W., Yu,S.H., Mulder,K.L. and Westphal,H. (1992) Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl Acad. Sci. USA*, **89**, 6232–6236.
33. Orban,P.C., Chui,D. and Marth,J.D (1992) Tissue- and site-specific DNA recombination in transgenic mice. *Proc. Natl Acad. Sci. USA*, **89**, 6861–6865.
34. Peitz,M., Pfannkuche,K., Rajewsky,K. and Edenhofer,F. (2002) Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. *Proc. Natl Acad. Sci. USA*, **99**, 4489–4494.
35. Green,M. and Loewenstein,P.M. (1988) Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell*, **55**, 1179–1188.
36. Frankel,A.D. and Pabo,C.O. (1988) Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, **55**, 1189–1193.
37. Ford,K.G., Darling,D., Souberbielle,B. and Farzaneh,F. (2000) Protein transduction: a new tool for the study of cellular ageing and senescence. *Mech. Ageing Dev.*, **121**, 113–121.
38. Nagahara,H., Vocero-Akbani,A.M., Snyder,E.L., Ho,A., Latham,D.G., Lissy,N.A., Becker-Hapak,M., Ezhevsky,S.A. and Dowdy,S.F. (1998) Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27^{Kip1} induces cell migration. *Nature Med.*, **4**, 1449–1452.
39. Schwarze,S.R., Hruska,K.A. and Dowdy,S.F. (2000) Protein transduction: unrestricted delivery into all cells? *Trends Cell Biol.*, **10**, 290–295.
40. 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.
41. Miyagishi,M. and Taira,K. (2003) Strategies for generation of an siRNA expression library directed against the human genome. *Oligonucleotides*, **13**, 325–333.
42. Miyagishi,M., Fujii,R., Hatta,M., Yoshida,E., Araya,N., Nagafuchi,A., Ishihara,S., Nakajima,T. and Fukamizu,A. (2000) Regulation of Lef-mediated transcription and p53-dependent pathway by associating β-catenin with CBP/p300. *J. Biol. Chem.*, **275**, 35170–35175.
43. Fritsch,L., Martinez,L.A., Sekhri,R., Naguibneva,I., Gerard,M., Vandromme,M., Schaeffer,L. and Harel-Bellan,A. (2004) Conditional gene knock-down by CRE-dependent short interfering RNAs. *EMBO Rep.*, **5**, 178–182.
44. Gu,H., Marth,J.D., Orban,P.C., Mossmann,H. and Rajewsky,K. (1994) Deletion of a DNA polymerase b gene segment in T cells using cell type-specific gene targeting. *Science*, **265**, 103–106.
45. Utomo,A.R.H., Nikitin,A.Y. and Lee,W.H. (1999) Temporal, spatial and cell type-specific control of Cre-mediated DNA recombination in transgenic mice. *Nat. Biotechnol.*, **17**, 1091–1096.
46. Kato,Y. and Taira,K. (2003) Expression of siRNA from a single transcript that includes multiple ribozymes in mammalian cells. *Oligonucleotides*, **13**, 335–344.
47. Weinberg,R.A. and Penman,S. (1968) Small molecular weight monodisperse nuclear RNA. *J. Mol. Biol.*, **38**, 289–304.
48. Metzger,D., Clifford,J., Chiba,H. and Chambon,P. (1995) Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc. Natl Acad. Sci. USA*, **92**, 6991–6995.
49. Kellendonk,C., Tronche,F., Monaghan,A.P., Angrand,P.O., Stewart,F. and Schütz,G. (1996) Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res.*, **24**, 1404–1411.
50. Kuhn,R., Schwenk,F., Aguet,M. and Rajewsky,K. (1995) Inducible gene targeting in mice. *Science*, **269**, 1427–1429
51. Farr,A. and Roman,A. (1992) A pitfall of using a second plasmid to determine transfection efficiency *Nucleic Acids Res.*, **20**, 920.

THE JOURNAL OF GENE MEDICINE
J Gene Med 2004; 6: 715–723.
Published online 8 March 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.556

RESEARCH ARTICLE

Optimization of an siRNA-expression system with an improved hairpin and its significant suppressive effects in mammalian cells

Makoto Miyagishi[1,2†]
Hidetoshi Sumimoto[3†]
Hiroyuki Miyoshi[4]
Yutaka Kawakami[3]
Kazunari Taira[1,2*]

[1]*Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Hongo, Tokyo 113-8656, Japan*

[2]*Gene Function Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Central 4, 1-1-1 Higashi, Tsukuba Science City 305-8562, Japan*

[3]*Division of Cellular Signaling, Institute for Advanced Medical Research, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan*

[4]*Subteam for Manipulation of Cell Fate, BioResource Center, RIKEN Tsukuba Institute, 3-1-1 Koyadai, Tsukuba, 305-0074, Japan*

*Correspondence to: Kazunari Taira, Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Hongo, Tokyo 113-8656, Japan.
E-mail: taira@chembio.t.u-tokyo.ac.jp

†These two authors contributed equally to this work.

Received: 16 September 2003
Revised: 1 December 2003
Accepted: 2 January 2004

Abstract

Background RNA interference (RNAi) is a phenomenon in which expression of an individual gene can be specifically silenced by introducing a double-stranded RNA, one complementary to the gene, into cells. This phenomenon can be observed in mammalian cells when small interfering RNAs (siRNAs) are used, and is receiving attention as the most powerful tool for reverse genetics in the post genome era. Several groups have developed vector-based siRNA-expression systems that can induce RNAi in living cells.

Methods We describe here a comparative analysis of various siRNA-expression systems, in which we examined the effects of stem length, loop sequence and insertion of mutation(s) and/or bulges in the stem sequence on silencing effects and on the stability of the vectors.

Results As a result of the comparative analysis, we determined the following optimized siRNA-expression system: U6 promoter-driven hairpin-type dsRNA with 21-nt stem length, three to four mutations in the sense strand only, and the optimized 9-nt loop sequence, derived from microRNA. Moreover, we demonstrate that the siRNA-expression system with a tetracycline-regulated U6 promoter(s) could have the potential to control RNAi in cells, and that the HIV vector-mediated transfer of an siRNA-expression cassette into cells resulted in efficient silencing of a target gene at a multiplicity of infection as low as five.

Conclusion The mutated hairpin siRNAs and their genetically stable coding vectors could be very useful for gene knockdown experiments, and could further benefit gene therapy using RNAi.

Keywords RNA interference; U6 promoter; HIV vector; gene silencing; tetracycline-regulated promoter

Introduction

RNA interference (RNAi) is a recently discovered 'gene-silencing' phenomenon that is induced by double-stranded RNA (dsRNA) [1]; reviewed in [2]. Although RNAi works well in various organisms, the silencing of specific genes by RNAi has proved difficult to detect in mammalian systems because of the dsRNA-dependent nonspecific inhibition of protein synthesis, which is part of the host's system for defense against viral infections. However, recent works by Tuschl and colleagues demonstrated that 21- or 22-nt RNAs with

2-nt 3′-overhangs can induce gene silencing without the nonspecific inhibition of translation in cultured mammalian cells [3,4]. Furthermore, in the past year, various groups, including our own, have developed systems for vector-mediated specific RNAi in mammalian cells [5–12]; reviewed in [13,14]. These vector systems use a pol III promoter, such as the U6, H1, and tRNA promoters, and the systems have been classified into two groups depending on whether the expressed RNAs are tandem-type or hairpin-type (Figure 1A). The tandem-type siRNA-expression vector was designed to include sense and antisense RNAs, transcribed by two independent pol III promoters, and subsequently the sense and antisense RNAs anneal and form siRNA duplexes. The hairpin-type siRNA-expression vector was designed for transcription of hairpin RNA, which is subsequently processed into siRNA duplexes by intrinsic RNase III. We have attempted a systematic comparison and optimization of these systems using a dual luciferase reporter system, and our observations have resulted in the identification of a novel vector system that is mutationally more stable and more efficient than others reported to date. Furthermore, we demonstrate the regulatable RNAi by using the tetracycline-regulated U6 promoter, and the lentivirus-mediated delivery of an siRNA-expression cassette into cells, which allows the efficient silencing of the expression of a target gene.

Materials and methods

Culture and transfections of cells and assays of the expression of reporter genes

HeLa S3 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Cells were plated at 3×10^4 cells per well in a 48-well plate one day before transfection. HEK293/TetR cells (Flp-In™ T-REx™ 293 cell line) were purchased from Invitrogen (Carlsbad, CA, USA). Transfections were performed with Lipofectamine™ 2000 reagent (Life Technologies, Rockville, MD, USA) using 300, 30, or 3 ng of siRNA-expression plasmid, 30 ng of firefly luciferase-expression plasmid (pGL3; Promega, Madison, WI, USA), and 10 ng of *Renilla* luciferase-expression plasmid (pRL-RSV [15]), as described by Promega. Luciferase activities were analyzed after 24 h with a dual luciferase system

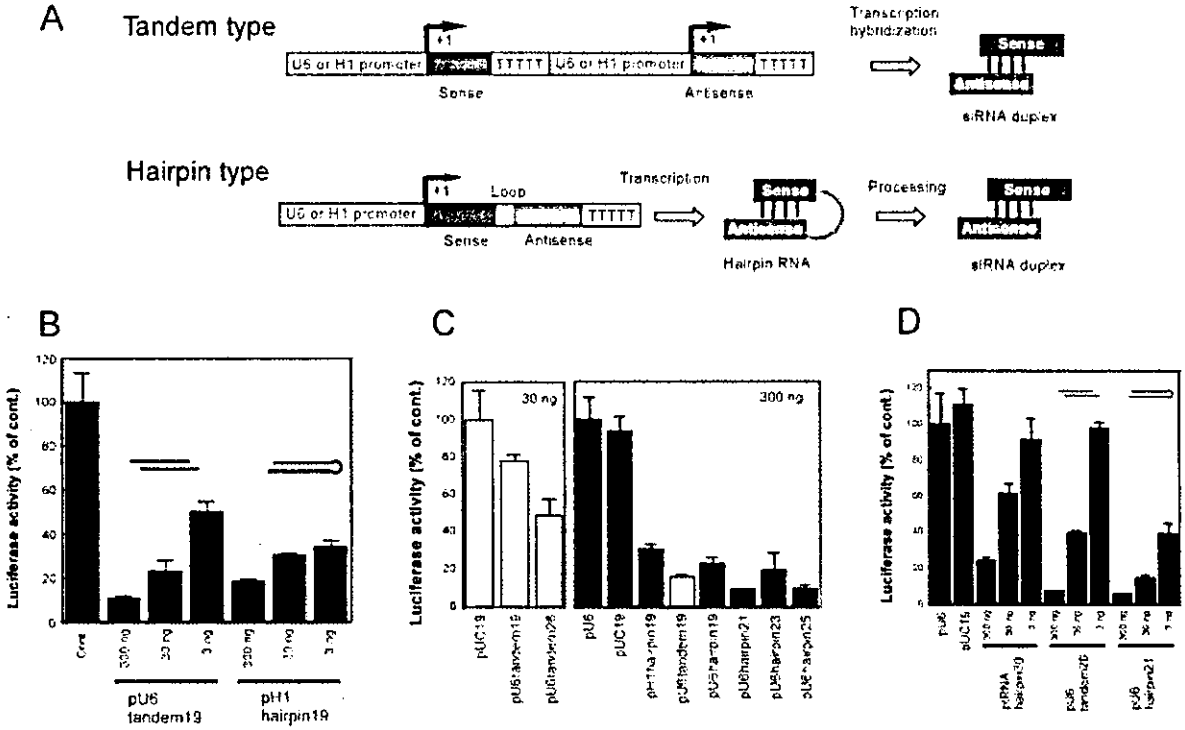

Figure 1. (A) Schematic representation of tandem-type and hairpin-type siRNA-expression vectors. In the tandem-type, both sense and antisense strands are driven by the U6 promoter. Subsequently, sense and antisense RNAs anneal and form siRNA duplexes with a 4-nt overhang at each 3′-end. In the hairpin-type, sense and antisense nucleotides are expressed as a single unit. The transcribed RNA forms a hairpin with a stem structure that is processed as indicated. (B) Comparative analysis of the effects of the U6 promoter-driven 19-nt tandem-type siRNA-expression vector and the H1 promoter-driven 19-nt hairpin-type siRNA-expression vector. 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 of triplicate assays. (C) Dependence of the effects of the stem length of hairpin-type and tandem-type siRNA-expression vectors on the RNAi activities. HeLa S3 cells were seeded in 48-well plates at a density of 3×10^4 per well and cotransfected with 10 ng of pRL-RSV, 30 ng of pGL3, and the indicated amount of each siRNA expression vector targeted to the firefly gene for luciferase. (D) Comparative analysis of optimized hairpin-type, tandem-type, and tRNA-type siRNA-expression vectors

(Promega). To ensure equal DNA amounts, empty plasmids were added at appropriate levels in each transfection.

Constructs used in transient transfections

We prepared tandem-type and hairpin-type siRNA-expression vectors using the pU6icassette vector [16], which contains a human U6 promoter and two BspMI sites. For the construction of tandem-type siRNA-expression plasmids, we amplified DNA fragments that included the sequences of the sense and antisense regions and the U6 promoter by PCR with the pU6icassette 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 BspMI, each fragment was ligated into the BspMI sites of the pU6icassette vector to yield a series of siRNA-expression vectors. For the 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 BspMI sites of the pU6icassette vector. H1 promoter-driven siRNA-expression vectors were generated as described by Brummelkamp *et al.* [5]. The tRNA promoter-driven siRNA-expression vectors were prepared by PCR-based subcloning into pUC-dt (a plasmid that contains the chemically synthesized promoter for a human gene for $tRNA^{Val}$ between the EcoRI and Sal I sites of pUC19 [17]). Sequences inserted immediately downstream of the U6 promoter were as follows (only the sense sequences are shown): in pU6tandem19, pU6hairpin19, and pH1hairpin19, GTG CGC TGC TGG TGC CAA C; in pU6tandem26, GTG CGC TGC TGG TGC CAA CCC TAT TC; in pU6hairpin21, GTG CGC TGC TGG TGC CAA CCC; in pU6hairpin23, GTG CGC TGC TGG TGC CAA CCC TA; in pU6hairpin25, GTG CGC TGC TGG TGC CAA CCC TAT T; and in ptRNAhairpin30, GTG CGC TGC TGG TGC CAA CCC TAT TCT CCT gaaa AGG AGA ATA GGG TTG GCA CCA GCA GCG CAC [both strands and the linker sequence (small letters) are shown for the tRNA-linked hairpin-type vector]. The sequences of the various mutated constructs are indicated in the figures. The miR30 microRNA precursor-directed siRNA-expression vectors (pri-miR30-directed siRNA-expression vectors) were generated by PCR-based subcloning into the pU6icassette and pEF9 [15] after digestion of products of PCR with BspMI and with BamHI plus EcoRI, respectively. The inserts of the constructs were generated by PCR with the following primers: 5′-ggc gga tcc ACC TGC cgg cca ccG tat att gct gtt gac agt gag cga cGT GCG tTG tTG GTG ttA AtC Ctt caa gag aGG-3′ and 5′-ggc gaa ttc ACC TGC tag cgc ata aaa att gaa gtc cga ggc agt agg cag cGT GCG CTG CTG GTG CCA ACC Ctc tct tga aGG-3′. The tetracycline-regulated tandem-type siRNA-expression vectors were generated from the pU6Teticassette vector as well as the tandem-type siRNA-expression vector. The pU6Teticassette vector was made from the pU6icassette by converting the wild-type U6 promoter into the tetracycline-regulated U6 promoter [18].

Construction of HIV vectors for generation of siRNA

HIV-U6

The human U6 promoter from pU6 was subcloned between the CTS (central termination sequence) and the CMV promoter of an HIV plasmid vector, CS-CDF-CG-PRE, which was described previously [19]. This vector is a self-inactivating (SIN) vector lacking its own promoter sequence in the 3′-LTR, which allows expression of the green fluorescence protein (GFP) under the control of the CMV promoter. It includes a central polypurine tract (cPPT) and the posttranscriptional regulatory element (PRE) of woodchuck hepatitis virus for enhanced nuclear transport of the transcribed mRNA. We made a series of siRNA-expression plasmids that were based on the HIV-U6 plasmid. The strategy for generating siRNA was described above. Products of PCR were inserted into two *Bsp*MI sites downstream of the U6 promoter.

HIV vectors

HIV vectors were produced by transfecting 293T cells by calcium phosphate transfection, as described previously, with HIV plasmid vectors, pMD.G (VSV-G env-expression plasmid), pMDLg/p.RRE (a third-generation packaging plasmid) and pRSV Rev (a Rev-expression plasmid). The latter two plasmids were provided by Cell Genesys (Foster City, CA, USA). Each vector-containing supernatant was collected, pooled, passed through a 0.45-μm filter concentrated with CENTRIPREP YM-50 (Millipore Co., Bedford, MA, USA), and stored at −80 °C until use. Titers of preparations of vectors were determined by infecting 293T cells. The transduction of genes to target cells was monitored in terms of expression of GFP, which was examined with a FACSCalibur™ flow cytometer and CellQuest™ software (Becton-Dickinson, San Jose, CA, USA). The titers of the concentrated preparations of vectors used in this study ranged from approximately 1×10^8 to 3×10^8 transducing units (TU)/ml.

Assays of expression of reporter genes after transfections with HIV vectors

293T cell were infected with either HIV-U6, HIV-tandem, or HIV-hairpin (C > T5) at an MOI of five. Each HIV vector directed the generation of siRNA against firefly luciferase mRNA. The sequences of each sense or antisense fragment are described above. Expression of GFP was examined 3 days after infection and 100% of infected cells were found to be GFP-positive by flow cytometry. The infected cells were transfected, using calcium phosphate, with 100 ng of the firefly luciferase expression vector (pGL3) and 100 ng of the *Renilla* luciferase expression

vector (pRL-SV40; Promega) in wells of 6-well plates. Luciferase activities were analyzed 24 h after the second transfection with the Dual-Glo™ luciferase assay system (Promega) and a Berthold luminometer. Each value for firefly luciferase was normalized by reference to the activity of *Renilla* luciferase activity. Assuming that the value for control 293T cells transfected with HIV-U6 was 100%, we calculated each value relative to the control value.

Results

Comparison of the suppressive effects of various siRNA-expression systems

To compare the suppressive effect of tandem-type and hairpin-type siRNA-expression vectors, we constructed two vectors that contained the tandem-type and the hairpin-type, respectively, of the siRNA-expression cassette, targeted to the same site in the transcript of the firefly gene for luciferase (Figure 1A). The hairpin-type siRNA-expression vector was designed to include a 19-nucleotide (nt) sense target sequence, followed by a 9-nt loop sequence, a 19-nt antisense target sequence and a pol III terminator, all located downstream of a promoter [5,20]. The tandem-type siRNA-expression vector was designed for transcription of 19-nt sense and antisense RNAs with a 3′ terminal polyuridine tail, from two independent U6 promoters [8,20].

We cotransfected HeLa S3 cells with the appropriate siRNA-expression vector, targeted to the transcript of the firefly gene for luciferase, a firefly luciferase-expression vector (pGL3) and a *Renilla* luciferase-expression vector (pRL-RSV), and then we measured luciferase activities 24 h after transfection. As shown in Figure 1B, there was no apparent difference in terms of the suppressive effect on the expression of firefly luciferase between the hairpin-type and the tandem-type siRNA-expression vectors. However, at lower concentrations of plasmids, the hairpin-type siRNA-expression vector appeared to have a slightly stronger suppressive effect than the tandem-type siRNA-expression vector.

To compare the two vector systems in further detail, we examined the dependence on the length of the sense and antisense sequences in each system. We constructed a tandem-type siRNA-expression vector, which encoded 26-nt sense and antisense RNAs, and compared its activity with that of the 19-nt dsRNA-expressing tandem-type vector. As shown in Figure 1C, the 26-nt sense and antisense-expressing tandem-type vector was a much stronger suppressor than the 19-nt sense and antisense-expressing tandem-type vector. We next made a series of U6-driven hairpin-type siRNA-expression vectors, which encoded 21-nt, 23-nt and 25-nt hairpin RNA, respectively, and evaluated their activities. The 21-nt and 25-nt hairpin RNAs were significantly more active suppressors of expression of the firefly luciferase gene than the 19-nt hairpin RNA. Since human Dicer processes dsRNA more efficiently at 20–21 nt in length from the terminus [21–23], the difference in suppressive activity between the 19-nt and the longer than 21-nt hairpin RNAs might reflect the efficiency of processing of hairpin RNA into siRNA by human Dicer.

We next compared the 21-nt hairpin-type vector and the 26-nt tandem-type vector, which had the highest suppressive activity of each respective vector system. At a high concentration of the siRNA-expression vector (300 ng), the hairpin-type and the tandem-type siRNA-expression vectors had similar strong suppressive activity. At a low concentration of the siRNA-expression vector, the hairpin-type siRNA-expression vector had significantly higher suppressive activity than the tandem-type siRNA-expression vector (Figure 1D).

Technical problems associated with the conventional hairpin-type siRNA-expression vectors and their solutions

When constructing hairpin-type siRNA-expression vectors, we faced two serious technical problems. First, it was difficult to sequence constructs that contained a hairpin region, probably because of the tight palindromic structure. Second, approximately 20–40% of constructs were mutated within the hairpin region of the constructs upon introduction into *E. coli*. Sequencing of mutated constructs revealed that constructs with two or more point mutations or insertions/deletions in the sense or antisense region could be sequenced without any problems. Therefore, we examined whether constructs with such mutations or insertions might retain silencing activity.

As many as seven point mutations from C to T, which generated G:U base-paring, in the sense strand did not affect the silencing effect (Figure 2). As many as four bulge insertions in the sense strand also did not affect the activity. By contrast, the constructs with five G to A mutations or more than four A to C mutations had significantly reduced silencing activity. In particular, insertion of a single G to A mutation in the antisense strand [G > A1(AS)], which produces a mismatch, resulted in reduced suppression, and insertions of two or five G to A mutations [G > A2(AS) or G > A5(AS)] in the antisense strand resulted in almost no silencing of the target gene. Therefore, it appears that, in practice, it is not possible to introduce mutations and bulges into the antisense strand and retain activity.

It should also be emphasized that, while generating these constructs with mutated hairpin structures, we found 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. This property of such constructs is obviously important for the maintenance of corresponding plasmids in their prokaryotic host. Collectively, our data suggested that the way to create a useful hairpin-type siRNA-expression vector was to introduce multiple C

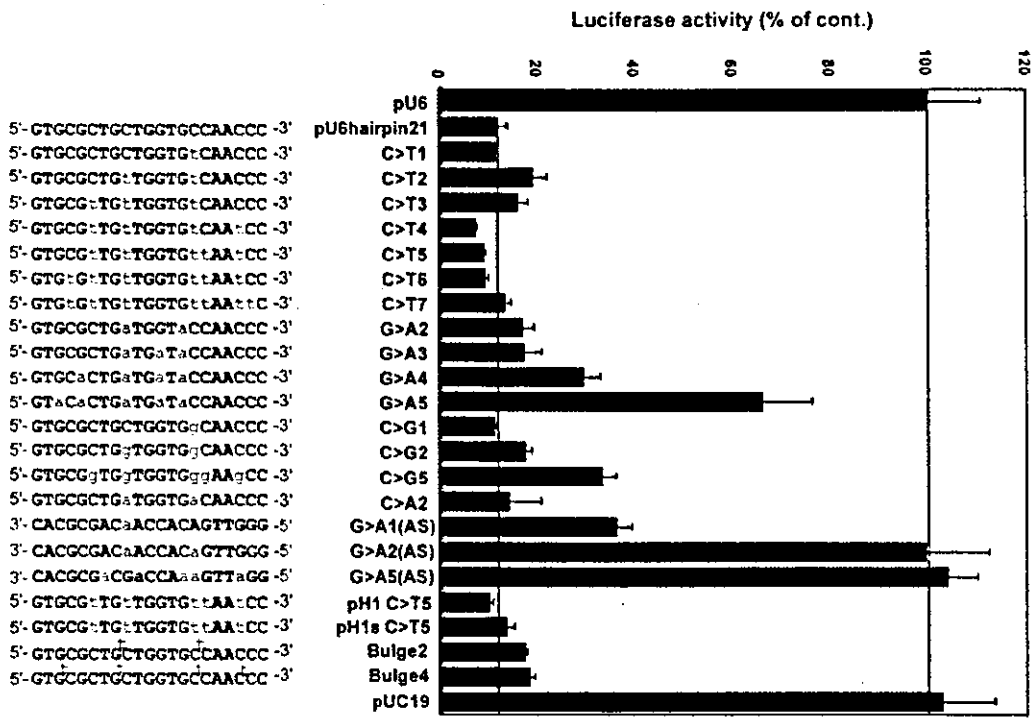

Figure 2. The effects of various mutations and/or bulge insertions in the hairpin sequence on silencing activity. HeLa S3 cells were seeded in 48-well plates at a density of 3×10^4 per well and cotransfected with 10 ng of pRL-RSV, 30 ng of pGL3, and 30 ng of an siRNA-expression vector with the indicated mutations, targeted to the firefly gene for luciferase. Letters in red indicate mutations or bulges

to T mutations, which not only prevented unwanted mutations, but also enhanced the silencing activity (compare results for pU6hairpin21 with C > T4, C > T5, or C > T6 in Figure 2, and that for pU6hairpin21 with C > T5 in Figure 3A for increased activities due to mutations) and/or to introduce multiple bulges into the sense region.

Effects of the loop sequence on silencing

We next examined the effects of the loop sequence on the silencing activity of hairpin-type siRNA-expression vectors. Brummelkamp *et al.* [5] reported that, when a 19-bp stem sequence was used, hairpin RNA with a 9-nt loop sequence had greater silencing activity than the corresponding RNA with a shorter loop sequence, and Yu *et al.* [12] used a 3-nt loop when they generated hairpin RNA with a 19-bp stem sequence. We wondered whether the natural loop sequence of microRNAs might be a preferable sequence for the hairpin RNA produced from hairpin-type siRNA-expression vectors and, therefore, we examined seven loop sequences derived from human microRNAs for analysis.

In transfection experiments using hairpin-type siRNA-expression vectors targeted to the firefly gene for luciferase, each included a microRNA-derived loop sequence, we found that all the hairpin-type siRNA-expression vectors that we tested had similar or even greater silencing activity when compared with the corresponding vectors with the 9-nt loop sequence (Figure 3A). In particular, loop2 and loop7 had slightly higher activity. Although it appeared that the loop sequence affected silencing activity only marginally (Figure 3A), in practice we found that, to achieve the suppressive activity of C > T5 loop7, several-fold higher concentrations of C > T5 with the original 9-nt loop were required (data not shown). Therefore, inclusion of a natural microRNA-derived loop was advantageous, especially when the copy number of the template was low.

Attempt to use pri-miRNA precursors to generate siRNAs

In a recent report, Y. Lee *et al.* [24] proposed that microRNAs are processed, via multiple steps, from long precursor transcripts, designated pri-miRNAs. This finding prompted us to attempt to generate siRNA using the precursor to microRNA. We anticipated efficient processing, transport to the cytoplasm, and incorporation into RISC by the endogenous processing system. We constructed expression vectors with the pol II and with the pol III promoter-driven precursor [25] to miR30 microRNA (pri-miR30), in which the sequence corresponding to the microRNA was replaced by the target sequence for the firefly gene for luciferase.

Using the U6 promoter-driven pri-miR30-expression vector targeted to firefly luciferase, we observed a 55% reduction in luciferase activity, but the effect was lower than that obtained with the hairpin-type siRNA-expression

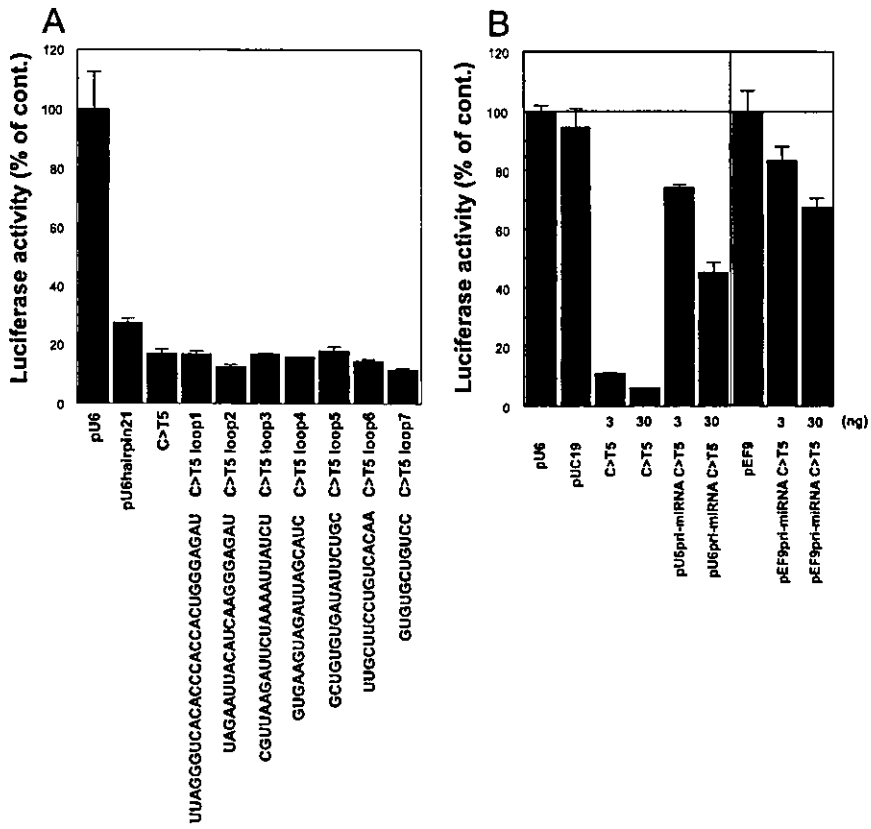

Figure 3. (A) The effects of the loop sequence on silencing. HeLa S3 cells were seeded in 48-well plates at a density of 3×10^4 per well and cotransfected with 10 ng of pRL-RSV, 30 ng of pGL3, and 30 ng of an siRNA-expression vector with the indicated loop sequence and the hairpin sequence with five C to T mutations, targeted to the firefly gene for luciferase. (B) The effects of the pol III promoter (U6 promoter) or pol II promoter (elongation factor 1 alpha promoter)-driven miR30 microRNA precursor expression vector

system, and the silencing activity of the pol II promoter-driven pri-miR30-expression vector was very much lower (Figure 3B). The siRNA-production system using the precursor to microRNA needs to be refined if we are to achieve efficient silencing.

Significant silencing by an HIV vector-derived optimized hairpin-expression system

Next, we attempted to use a lentivirus vector to deliver our optimized siRNA-expression cassette into cells. We introduced five C to T mutations into the sense strand of the hairpin RNA and postulated that our HIV-mediated delivery system would allow very efficient transfection of nondividing target cells, such as primary neurons. We prepared a lentivirus that included a hairpin-type or a tandem-type siRNA-expression cassette (the tandem-type has no mutation in the sense strand) that was targeted to the firefly gene for luciferase and transfected 293T cells (Figure 4). One day after this first transfection, we transfected the same cells with a firefly luciferase-expression vector and a *Renilla* luciferase-expression vector and examined the expression of the two luciferases. As shown in Figure 4B, the firefly luciferase activity in cells that had been transfected with the hairpin-type siRNA-expression lentivirus vector was approximately 90% lower than that in mock-transfected cells. Figures 4C and 4D demonstrate the transfection efficiencies of luciferase expression vectors. The tandem-type siRNA-expression cassette had no effect on the relative luciferase activity, and all cells transfected with an siRNA-expressing lentivirus vector showed no evidence of any toxic effects (Figures 4B and 4C).

This above result was consistent with the results of transient expression assays in which a hairpin-type siRNA-expression vector had a higher suppressive effect on luciferase activity than the tandem-type expression vector at a low concentration of plasmid. The correspondence reflects the observation that siRNA-expression cassettes are integrated into the genome in low numbers under our transfection conditions. However, it should be mentioned that tandem constructs in lentiviral vectors could be used under the condition in that titer of virus is sufficiently high and/or the target site is favorable [26,27].

Regulatable RNAi by using a tetracycline-controlled U6 promoter

In view of the potential utility of the siRNA expression vector and with the goal of analyzing the functions of

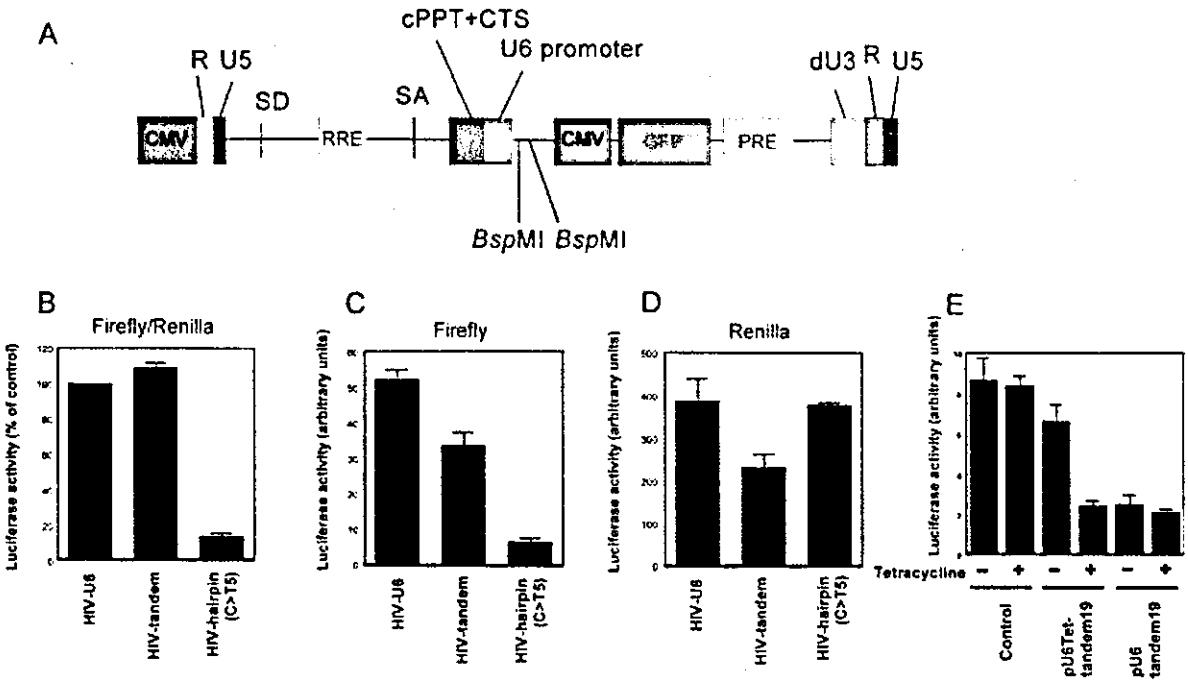

Figure 4. (A) Schematic representation of the HIV vector construct. HIV-U6 is the basic HIV vector plasmid and it includes an internal cassette for expression of siRNA. Two *Bsp*MI sites downstream of the U6 promoter are cloning sites for sense and antisense nucleotides. CMV, cytomegalovirus promoter; SD, splice donor site; SA, splice acceptor site; RRE, rev-responsive element; cPPT, central polypurine tract; CTS, central termination sequence; GFP, cDNA for green fluorescence protein; and PRE, posttranscriptional regulatory element of woodchuck hepatitis virus. (B) Effects of HIV siRNA-expression vectors on luciferase activity. Each bar indicates an average value and vertical bars indicate standard errors of triplicate assays. These results are representative of the results of four individual experiments that gave similar results. (C, D) Activities of firefly and *Renilla* luciferases, in arbitrary units, from the indicated transfectants. All cells in each cell line were positive for expression of GFP (data not shown). Each bar indicates an average value and vertical bars indicate standard errors of triplicate assays. (E) Control of RNAi by the tetracycline-regulated U6 promoter. HEK293/TetR cells were seeded in 48-well plates at a density of 3×10^4 cells/well and cotransfected with 10 ng of pRL-RSV, 30 ng of pGL3, and 300 ng of wild type or tetracycline-regulated siRNA expression vector (pU6Tet-tandem19) targeted to the firefly gene for luciferase, or an empty vector in the presence or absence of tetracycline (5 mg/ml). Twenty-four hours after transfection, the luciferase activities were determined

lethal genes in the future, we next tried to develop a system for the inducible expression of RNAi by using the tetracycline-regulated U6 promoter, which has previously been characterized in our laboratory [18]. We constructed a firefly luciferase-targeted tandem-type siRNA expression vector under the control of the tetracycline-regulated U6 promoters. HEK293/TetR cells, which stably express the gene for tetracycline repressor, were cotransfected with the tetracycline-controlled tandem-type siRNA-expression vector, the firefly luciferase-expression vector, and the *Renilla* luciferase-expression vector in the presence or absence of tetracycline. The suppressive activity of the wild-type siRNA-expression vector was unaffected by the presence or absence of tetracycline (Figure 4E). In contrast, the construct with the tetracycline-regulated U6 promoter clearly inhibited the expression of firefly luciferase in the presence of tetracycline, but did not suppress the gene expression in the absence of tetracycline (Figure 4E). This result demonstrates that the siRNA-expression system with the tetracycline-regulated U6 promoter(s) does indeed have the potential to control RNAi in cells.

Discussion

In this study, we compared the effects of various siRNA-expression systems and optimized some aspects of these systems by introducing multiple mutations and/or bulges into the sense strand of the hairpin RNA. Our optimized siRNA-expression system is summarized in Figure 5. The system employs U6 promoter-driven hairpin-type dsRNA with 21-nt stem length, three to four mutations in the sense strand only, and the optimized 11-nt loop sequence, derived from microRNA (miR-26b).

In the experiments to compare RNAi activity between the hairpin-type and the tandem-type siRNA-expression vectors, the hairpin-type expression vector showed higher suppressive activity than the tandem-type siRNA-expression vector at lower concentrations of plasmid. There are at least two possible interpretations of these observations. First, since the sense and antisense RNAs of hairpin RNA encoded by the hairpin-type siRNA-expression vector lie "*in cis*", the efficiency of annealing of the sense to the antisense RNAs to form a hairpin RNA might be much higher than that of separately expressed

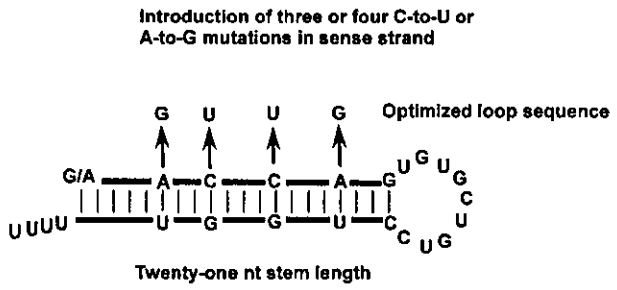

Figure 5. Schematic representation of the proposed siRNA-expression system

sense and antisense RNAs as encoded by the tandem-type siRNA-expression vector, in particular at a low concentration [28]. Thus, the double-stranded RNAs that are formed with the greater efficiency can be protected by proteins, such as RISC, from attacks by ribonucleases [8]. Second, exportin 5 was recently identified as an export factor for microRNA precursors [29] that form hairpin structures and, thus, it is possible that hairpin RNA, expressed from the hairpin-type siRNA-expression vector, might be a favorable substrate for exportin 5 and might be transported more efficiently from the nucleus to the cytoplasm than annealed siRNA expressed from a tandem-type siRNA-expression vector. However, since the transport of tRNA-based, hairpin RNA-derived siRNAs appears not to depend on such an export factor [6], and since the activity of the tRNA-based siRNAs is similar to that of separately expressed sense and antisense RNAs, at least with respect to the present specific target site (Figure 1D), enhanced transport from the nucleus to the cytoplasm is unlikely to be, by itself, the sole cause of the enhanced activity of the U6 promoter-driven hairpin RNAs.

We also showed that HIV vector-mediated transfer of an siRNA-expression cassette resulted in efficient silencing of a target gene, providing one of the first examples of the use of a lentivirus vector for the introduction of an siRNA-expression cassette into mammalian cells [20,30–33].

It was reported recently that adenovirus and retrovirus systems can be used for the delivery of siRNA-expression cassettes into mammalian cells [34–36]. However, large numbers of copies of an integrated siRNA-expression cassette appeared to be needed in the cited studies for adequate silencing of target genes. Our transfection experiments with multiplicities of infection (MOIs) of lentivirus as low as 5, in which only small numbers of copies of siRNA-expression cassettes could be integrated into the genome, revealed strong and maybe genetically stable suppressive effects on the expression of the target gene. The difference between earlier results and ours might be due to our use of an optimized expression system and/or a favored target site.

One question arises from our experiments: how is it that the lentivirus RNA genome, which includes the target sequence of the siRNA in a hairpin sequence during the production of RNA viruses, does not become a target for siRNA? The simplest explanation might be the dependence on target site of the silencing effect. The target site in the lentivirus genome probably forms a tight secondary structure because of the hairpin sequence. Then the hairpin target would not be accessible to siRNAs and would, as a consequence, escape cleavage by RISC [37].

In conclusion, we believe that our observations should be useful for further attempts at the silencing of the expression of specific genes and in gene therapies that exploit RNAi, especially because we can now predict the best target sites with the most optimized parameters, resulting in the correlation coefficient of up to 0.8 for predicted vs. actual results, in consideration of off-target effects and single nucleotide polymorphisms (SNPs) [28,38,39]. Moreover, functionally important microRNA can also be controlled by the present vectors [40,41].

References

1. Fire A, Xu S, Montgomery MK, *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; **391**: 806–811.
2. Zamore PD. RNA interference: listening to the sound of silence. *Nat Struct Biol* 2001; **8**: 746–750.
3. Caplen NJ, Parrish S, Imani F, *et al.* Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci U S A* 2001; **98**: 9742–9747.
4. Elbashir SM, Harborth J, Lendeckel W, *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001; **411**: 494–498.
5. Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 2002; **296**: 550–553.
6. Kawasaki H, Taira K. Short hairpin type of dsRNAs that are controlled by $tRNA^{Val}$ promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* 2003; **31**: 700–707.
7. Lee NS, Dohjima T, Bauer G, *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* 2002; **20**: 500–505.
8. Miyagishi M, Taira K. U6 promoter-driven siRNAs with four-uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat Biotechnol* 2002; **20**: 497–500.
9. Paddison PJ, Caudy AA, Bernstein E, *et al.* Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 2002; **16**: 948–958.
10. Sui G, Soohoo C, Affar el B, *et al.* A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc Natl Acad Sci U S A* 2002; **99**: 5515–5520.
11. Paul CP, Good PD, Winer I, *et al.* Effective expression of small interfering RNA in human cells. *Nat Biotechnol* 2002; **20**: 505–508.
12. Yu JY, DeRuiter SL, Turner DL. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc Natl Acad Sci U S A* 2002; **99**: 6047–6052.
13. McManus MT, Sharp PA. Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* 2002; **3**: 737–747.
14. Tuschl T. Expanding small RNA interference. *Nat Biotechnol* 2002; **20**: 446–448.
15. Miyagishi M, Fujii R, Hatta M, *et al.* Regulation of Lef-mediated transcription and p53-dependent pathway by associating β-catenin with CBP/p300. *J Biol Chem* 2001; **275**: 35 170–35 175.
16. Miyagishi M, Taira K. Expression of siRNA from a pol III promoter in mammalian cells. In *Perspectives in Gene Expression*, Appasani K (ed). The Eaton Publishers: Westboro, MA, 2003; 361–376.
17. Koseki S, Tanabe T, Tani K, *et al.* Factors governing the activity *in vivo* of ribozymes transcribed by RNA polymerase III. *J Virol* 1999; **73**: 1868–1877.

18. Ohkawa J, Taira K. Control of the functional activity of an antisense RNA by a tetracycline-responsive derivative of the human U6 snRNA promoter. *Hum Gene Ther* 2000; **11**: 577–585.
19. Sumimoto H, Tsuji T, Miyoshi H, *et al.* Rapid and efficient generation of lentivirally gene-modified dendritic cells from DC progenitors with bone marrow stromal cells. *J Immunol Methods* 2002; **271**: 153–165.
20. Taira K, Miyagishi M. siRNA expression systems and methods for producing functional gene knock-down cells. Japanese Patent Application, 2001; 2001–363385 (PCT/JP02/12447).
21. Provost P, Dishart D, Doucet J, *et al.* Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J* 2002; **21**: 5864–5874.
22. Zhang H, Kolb FA, Brondani V, *et al.* Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J* 2002; **21**: 5875–5885.
23. Kawasaki H, Suyama E, Iyo M, *et al.* Recombinant human Dicer (re-hDicer) generated siRNAs significantly induce a specific but target site-independent gene silencing in human cells. *Nucleic Acids Res* 2003; **31**: 981–987.
24. Lee Y, Jeon K, Lee JT, *et al.* MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* 2002; **21**: 4663–4670.
25. Kato Y, Taira K. Expression of siRNA from a single transcript that includes multiple ribozymes in mammalian cells. *Oligonucleotides* 2003; **13**: 335–343.
26. Banerjea A, Li MJ, Bauer G, *et al.* Inhibition of HIV-1 by lentiviral vector-transduced siRNAs in T lymphocytes differentiated in SCID-hu mice and CD34+ progenitor cell-derived macrophages. *Mol Ther* 2003; **8**: 62–71.
27. Li MJ, Bauer G, Michienzi A, *et al.* Inhibition of HIV-1 infection by lentiviral vectors expressing Pol III-promoted anti-HIV RNAs. *Mol Ther* 2003; **8**: 196–206.
28. Miyagishi M, Taira K. Strategies for generation of an siRNA expression library directed against the human genome. *Oligonucleotides* 2003; **13**: 325–333.
29. Lund E, Guttinger S, Calado A, *et al.* Nuclear export of microRNA precursors. *Science* 2003; 10.1126 science.1090599 (Science Express Reports).
30. Abbas-Terki T, Blanco-Bose W, Deglon N, *et al.* Lentiviral-mediated RNA interference. *Hum Gene Ther* 2002; **13**: 2197–2201.
31. Qin XF, An DS, Chen IS, *et al.* Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. *Proc Natl Acad Sci U S A* 2002; **100**: 183–188.
32. Tiscornia G, Singer O, Ikawa M, *et al.* A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. *Proc Natl Acad Sci U S A* 2003; **100**: 1844–1848.
33. Rubinson DA, Dillon CP, Kwiatkowski AV, *et al.* A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat Genet* 2003; **33**: 401–406.
34. Barton GM, Medzhitov R. Retroviral delivery of small interfering RNA into primary cells. *Proc Natl Acad Sci U S A* 2002; **99**: 14943–14945.
35. Brummelkamp TR, Bernards R, Agami R. Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* 2002; **2**: 243–247.
36. Xia H, Mao Q, Paulson HL, *et al.* siRNA-mediated gene silencing in vitro and in vivo. *Nat Biotechnol* 2002; **20**: 1006–1010.
37. Yoshinari K, Miyagishi M, Taira K. Effects on RNAi of the structure, sequence and position of the targeted region. *Nucleic Acids Res.* 2003; **32**: 291–299.
38. Taira K, Miyagishi M. Methods and devices for the prediction of favorable target sites of siRNA. *Japanese Patent Application* 2003–349 283.
39. http://www.igene-therapeutics.co.jp.
40. Kawasaki H, Wadhwa R, Taira K. World of small RNAs: from ribozymes to siRNA and miRNA. *Differentiation* 2004; in press.
41. Kuwabara T, Hsieh J, Nakashima K, *et al.* A small modulatory dsRNA specifies the fate of neural stem cells. *Cell* 2004; in press.

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Virus Research 102 (2004) 117–124
Virus Research
www.elsevier.com/locate/virusres

Generation of an shRNAi expression library against the whole human transcripts

Makoto Miyagishi [a,b], Sahohime Matsumoto [a,b,c], Kazunari Taira [a,b,*]

[a] *Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Hongo, Tokyo 113-8656, Japan*
[b] *Gene Function Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Central 4, 1-1-1 Higashi, Tsukuba Science City 305-8562, Japan*
[c] *Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Hongo, Tokyo 113-0033, Japan*

Abstract

RNA interference is a phenomenon in which expression of an individual gene can be specifically silenced by introducing a double-stranded RNA, homologous to the gene, and is receiving attention as a powerful tool for reverse genetics in the post-genome era. Throughout our current research to generate an siRNA expression library for the whole human genome, we face many technical difficulties. We present here the strategies for overcoming some of the difficulties, including the development of genetically stable and highly active siRNA expression vectors, the selection procedure of the favorable target sites, and the efficient and low cost procedure for constructing an siRNA expression library. Furthermore, we demonstrate that the screening using the constructed siRNA-expression library indeed works, by evaluating siRNA-expression library against apoptosis-related genes.

Keywords: RNA interference; U6 promoter; Algorithm; RNAi library

1. Introduction

Recent advances in the Human Genome Project and availability of the complete genome sequence have clarified the existence of numerous genes whose functions are unknown. Together with the hybrid ribozyme technology (Warashina et al., 2001; Kawasaki and Taira, 2002; Kawasaki et al., 2002; Suyama et al., 2003a,b), RNA interference (RNAi) is receiving attention as the technology that could unveil their functions (McManus and Sharp, 2002).

RNAi is an evolutionarily conserved process in plants and animals by which double-stranded RNA (dsRNA) induces sequence specific degradation of homologous RNA (Fire et al., 1998; reviewed in Zamore, 2001). Although RNAi works well in various organisms, the silencing of specific genes by RNAi has proven difficult to detect in mammalian systems, because of the dsRNA-dependent nonspecific inhibition of protein synthesis by the PKR pathway and the nonspecific RNA degradation by activation of RNase L (Elbashir et al., 2001). However, recent works by Tuschl and co-workers demonstrated that 21- or 22-nt RNAs with two nt 3′-overhangs (siRNA) can induce gene silencing without the nonspecific inhibition of the gene expression in cultured mammalian cells (Caplen et al., 2001; Elbashir et al., 2001). Furthermore, in the past year, 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; reviewed in McManus and Sharp, 2002; Tuschl, 2002). These vector systems use a pol III promoter, such as the U6, H1, and tRNA promoters, and the systems have been classified into two groups depending on whether the expressed RNAs are tandem-type or hairpin-type.

Since gene silencing by siRNA is highly specific and has extremely high activity, genome-wide and comprehensive gene analysis in various organism should be expected. Recently, in *Caenorhabditis elegans*, a systematic functional analysis of a number of genes by RNAi has been performed, and many of these genes were identified as genes that show mutant phenotype (Fraser et al., 2000; Gonczy et al., 2000). In the near future, it is predicted that a similar genome-wide RNAi library approach could be extended

* Corresponding author. Tel.: +81-3-5841-8828/298-61-3015; fax: +81-3-5841-7340/8828.
E-mail address: taira@chembio.t.u-tokyo.ac.jp (K. Taira).

0168-1702/$ – see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.virusres.2004.01.022

to mammalian cells by an siRNA oligonucleotide library or an siRNA expression library, with the success that randomized ribozyme libraries have had in the 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). In fact, several groups, including ours, have already begun the generation of an siRNA library against the whole human genome. In this report, we present the practical problems and methods for solving some of the difficulties that were faced in generating the siRNA expression library.

2. Materials and methods

2.1. Constructs used in transient transfections

We prepared tandem-type and hairpin-type siRNA expression vectors using the piGENE hU6 vector (Miyagishi and Taira, 2003; iGENE Therapeutics, Inc.; http//www.iGENE-therapeutics.co.jp), which contains a human U6 promoter and two *Bsp*MI sites. For the construction of tandem-type siRNA expression plasmids, we amplified DNA fragments that included the sequences of the sense and antisense regions and the U6 promoter by PCR with the piGENE hU6 vector as a template, and primers that included the sequences of the sense or antisense sequence plus the terminator (Fig. 1). 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 the 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 pU6tandem19, pU6hairpin19, and pH1hairpin19, GTG CGC TGC TGG TGC CAA C; in pU6tandem26, GTG CGC TGC TGG TGC CAA CCC TAT TC; in pU6hairpin21, GTG CGC TGC TGG TGC CAA CCC; in pU6hairpin23, GTG CGC TGC TGG TGC CAA CCC TA; in pU6hairpin25, GTG CGC TGC TGG TGC CAA CCC TAT T. The sequences of the various mutated constructs are indicated in Fig. 2.

2.2. Culture and transfection of cells and assays of the expression of reporter genes

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

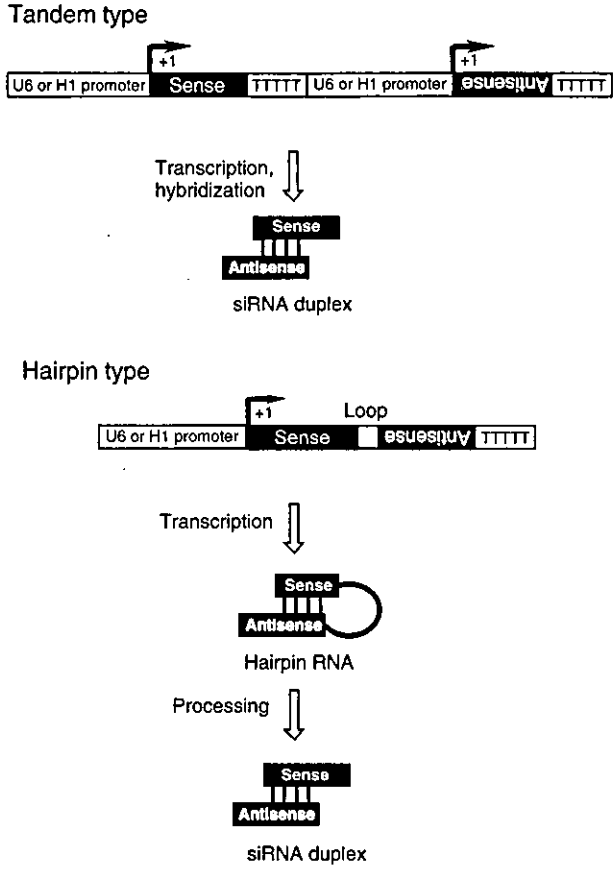

Fig. 1. Schematic representation of tandem-type and hairpin-type siRNA-expression vectors.

2.3. Algorithm for searching favorable target sites

Prediction modeling of the RNAi effect was performed by partial least squares (PLS) method. One hundred and fourteen data sets of siRNA sequences and their suppression values, obtained from published results and our own data, were analyzed and significantly correlated factors, including GC contents, were extracted. A PLS calibration model was made with the target site sequences, their suppression values and the extracted factors using the 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 were established, the suppression values at the various sites of the firefly luciferase were predicted using their target sequences.

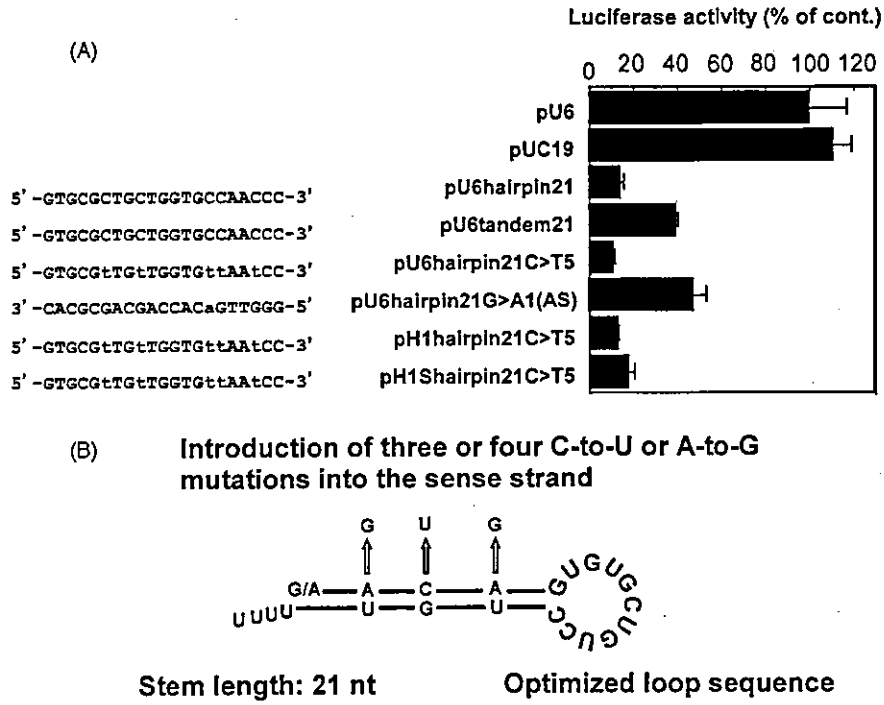

Fig. 2. (A) Comparative analysis of the effects of the U6 promoter-driven 21-nt tandem-type and 21-nt hairpin-type siRNA expression vectors, the H1 promoter-driven 21-nt hairpin-type siRNA-expression vector and the U6-promoter-driven 21-nt hairpin-type with various mutations in hairpin region. HeLa S3 cells were cotransfected with 10 ng of pRL-RSV, 30 ng of pGL3, and 30 ng of each siRNA expression vector targeted to the firefly gene for luciferase. The activities of firefly luciferase were normalized by reference to those of *Renilla* luciferase. Each bar indicates an average value and horizontal bars indicate standard errors of triplicate assays. (B) Schematic representation of the proposed siRNA expression system.

3. Results and discussion

3.1. Optimization of siRNA expression system: tandem-type or hairpin-type?

To construct the siRNA expression library, an optimization of the siRNA expression system, that has high suppressive activity and high genetic stability, should be initially addressed.

The major class of siRNA expression systems uses pol III promoter, such as U6 or H1 promoter. The systems have been classified into two groups, tandem-type or hairpin-type, depending on their expressing RNA (Fig. 1). In the case of tandem-type siRNA expression vector, both sense and antisense strands are driven separately by a respective pol III promoter, anneal in the cells and form siRNA duplexes with about four-nt overhang at each 3′-end (Lee et al., 2002; Miyagishi and Taira, 2002). In the hairpin-type of siRNA expression, sense and antisense nucleotides are connected by a loop and expressed as a single chain (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 will rapidly form a hairpin structure with a stem and loop and seems to be processed to siRNA by Dicer, like micro RNAs (miRNAs) (Ketting et al., 2001). We tried to examine which vector system had higher suppressive activity. We constructed both types of siRNA expression vectors, targeted against the same site of firefly luciferase gene, and compared their suppressive activities. Although both had high levels of suppressive activities at a high concentration of siRNA expression vectors, it was found that the hairpin-type siRNA expression vector (21-nt stem length) had significantly higher suppressive activity than the tandem-type siRNA expression vector at a low concentration of the siRNA expression vector (Fig. 2A).

3.2. Genetically stable hairpin-type siRNA expression vector

Although the hairpin-type siRNA expression vector has higher suppressive activity than the tandem-type siRNA expression vector, the palindromic sequence of the hairpin-type siRNA expression vector frequently causes two serious technical problems. Firstly, it is difficult to sequence constructs that contained a hairpin region, probably because of the tight palindromic structure. Secondly, approximately 20–40% of constructs are mutated within the hairpin region of the constructs upon introduction into *Escherichia coli*, which is the most serious problem for the library construction. We were aware that constructs with two or more point mutations or insertions/deletions in the sense or antisense region could be sequenced without any problems and some of them retained the suppressive activity. Therefore, we examined in detail the activities of constructs with various mutations and/or

insertions/deletions in the sense or antisense region (detailed analysis will be published elsewhere). As shown in Fig. 2A, as many as five-point mutations from C to T, which generated G:U base-pairing, in the sense strand did not affect the silencing effect (pU6hairpin21C > T5 in Fig. 2A). In contrast, insertion of a single C–T mutation in the antisense strand (pU6hairpinG > A1(AS) in Fig. 2A), which recognizes the target transcript, resulted in reduced suppression. Therefore, it appears that, in practice, it is not possible to introduce mutations into the antisense strand and retain activity.

While generating these constructs with mutated hairpin structures, we found 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). This property of such constructs is obviously important for the maintenance of corresponding plasmids in their prokaryotic host. Collectively, our data suggest that the way to generate a useful hairpin-type siRNA expression vector that enables us to construct an siRNA expression library is to introduce multiple C–T (or A–G) mutations only into the sense strand part of the hairpin (Fig. 2B).

3.3. Comparative analysis of various pol III promoters for silencing activity

We also compared the RNAi activity of the U6 promoter and the H1 promoter, both of which belong to the type II Polymerase III promoters, and contain the same consensus motifs, distal sequence element (DSE), proximal sequence element (PSE) and TATA box. As shown in Fig. 2A, H1 promoter-driven siRNA construct targeted against firefly luciferase has an activity comparable, but slightly lower, than that of the U6-driven siRNA construct (comparing pH1hairpin21C > T5 with pU6hairpin21C > T5 in Fig 2A, and other unpublished results). Therefore, we selected U6-driven siRNA expression vector as the promoter for the siRNA expression library. The tRNA-linked siRNA expression vector is a recently developed expression system, which would enable us to forcibly transport the hairpin RNA to the cytoplasm (Kawasaki and Taira, 2003). However, it requires a longer stem length (30-nt) and it would be more costly; hence we adopted the U6-driven siRNA expression vector for the construction of the siRNA expression library.

3.4. Optimization of the loop sequence

Different loop sequences of a hairpin-type siRNA expression were used according to respective researchers. For example, Agami's group uses a 9-nt loop sequence, 5′-UUCAAGAGA-3′, (Brummelkamp et al., 2002) and Paul et al. (2002) use a 4-nt, 5′-UUCG-3′, tetra nucleotide sequence. Our experiments using hairpin-type siRNA expression vector with randomized 6-nt loop sequences revealed that the hairpin RNAs with different loop sequences had different suppression activities (Kawasaki and Taira, 2003; Miyagishi et al., 2004), therefore, the loop sequence appears to somewhat influence the RNAi effect.

We wondered whether the natural loop sequence of miRNAs might be a preferable sequence for the loop of the hairpin RNA produced from hairpin-type siRNA expression vectors. Examination of several miRNA-derived loop sequences revealed that loop 7, which is 11-nt in length and

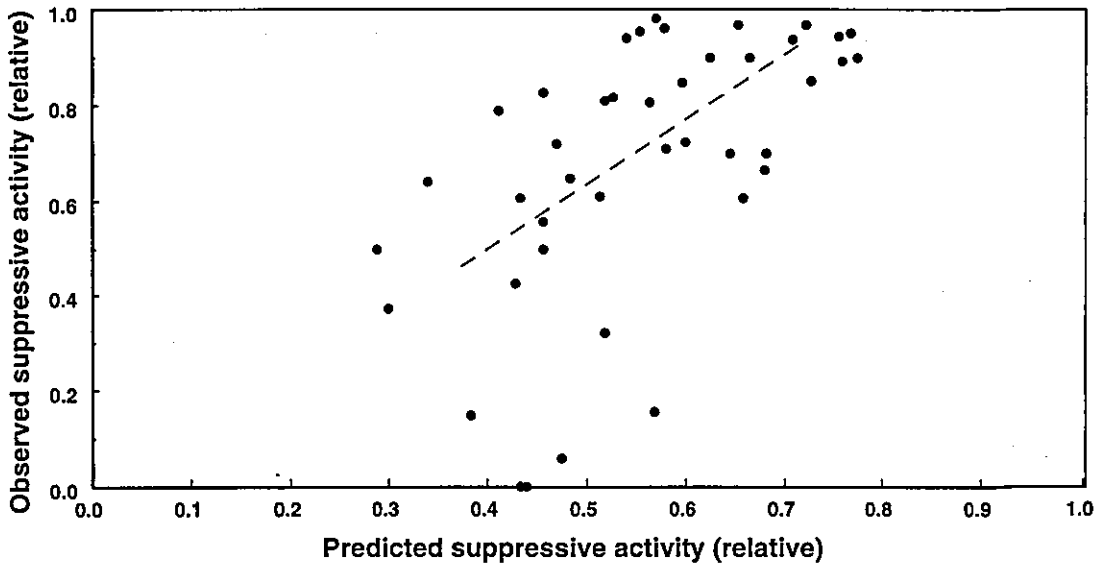

Fig. 3. Relationship between predicted and actual values of suppression activity. The actual values were measured by co-transfection experiments with 10 ng of pRL-RSV, 30 ng of pGL3, and 10 nM of each siRNA oligonucleotides targeted to the firefly gene for luciferase into HeLa cells. The activities of firefly luciferase were normalized by reference to those of *Renilla* luciferase and the suppression values were calculated as the ratio to the negative control. The predicted values were calculated from the target sequence of the siRNA with the PLS algorithm.

derived from the human miRNA miR-26b (the sequence is 5′-GUG UGC UGU CC-3′), is highly effective (the loop is shown in Fig. 2B). Therefore, we adopted this loop as the optimized loop sequence of the hairpin-type siRNA expression vector for the siRNA expression library.

3.5. Method for selection of favorable target site

One of the most important and critical problems of the siRNA expression library would be selection of the target sites. It is well-known that the effectiveness of siRNA is very dependent on the target site of a message and the possibility that siRNAs at an arbitrarily selected target site have high suppressive activity, seems to be only at about 20–40% or less. To generate an siRNA expression library of high-quality, an algorithm that can accurately predict an effective site must be developed. We analyzed siRNA data sets of several hundreds of target sites, including published work and our own experimental data, extracted several correlated factors, and generated an algorithm by nonlinear regression methods.

Fig. 3 shows predicted versus measured results for siRNAs targeted against firefly luciferase, which were *not* used in the generation of the algorithm. The correlation coefficient was not high (0.58) but, notably, all of the siRNAs that were calculated as scoring more than 0.73 showed high suppressive activities (more than 90%), therefore, especially at high prediction values, the predicted scores seem to be accurate. We use target sites that have a predicted suppressive activity of more than 0.75 for the construction of the siRNA expression library.

3.6. Confirmation of specificity of the target sequence

When one selects a certain target site in a specific gene, there is a need to check the specificity of the sequence of the

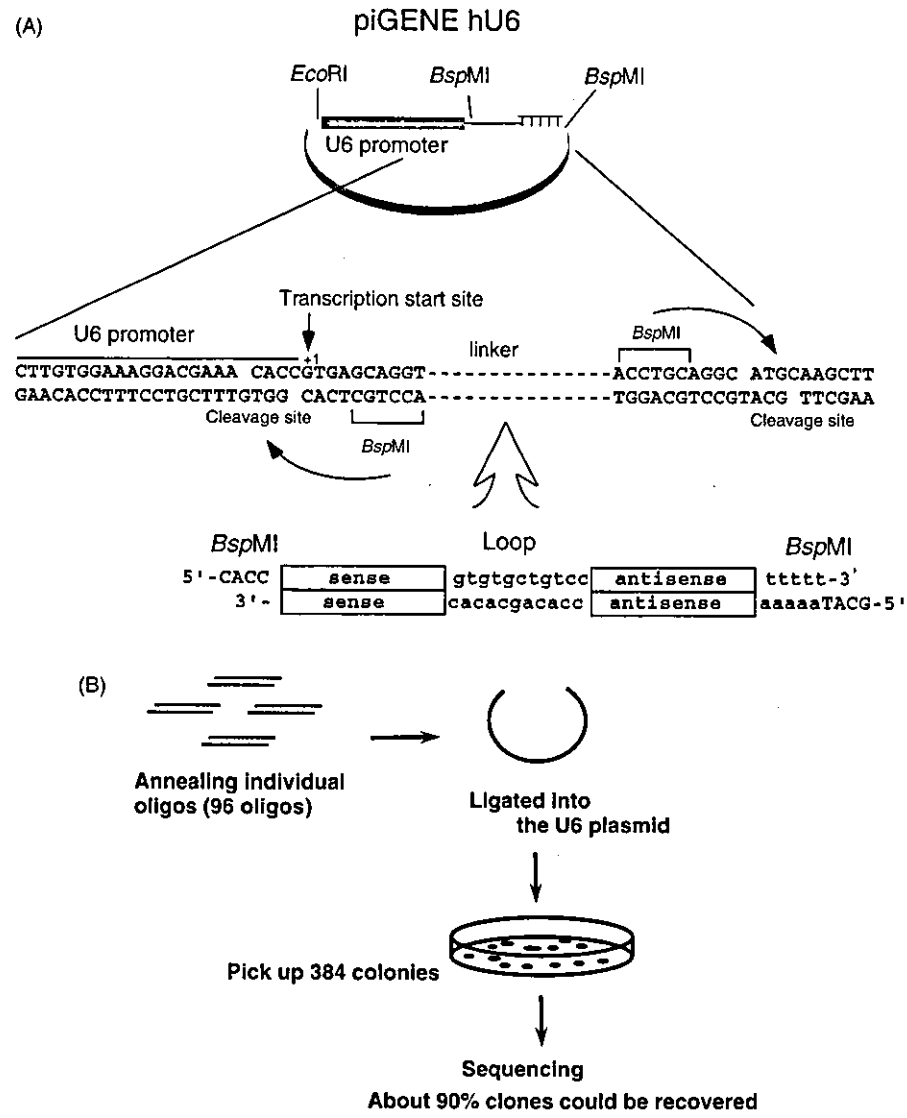

Fig. 4. (A) Cloning site and the insert sequence of hairpin-type siRNA-expression vector. (B) Our strategy for the construction of siRNA-expression library.

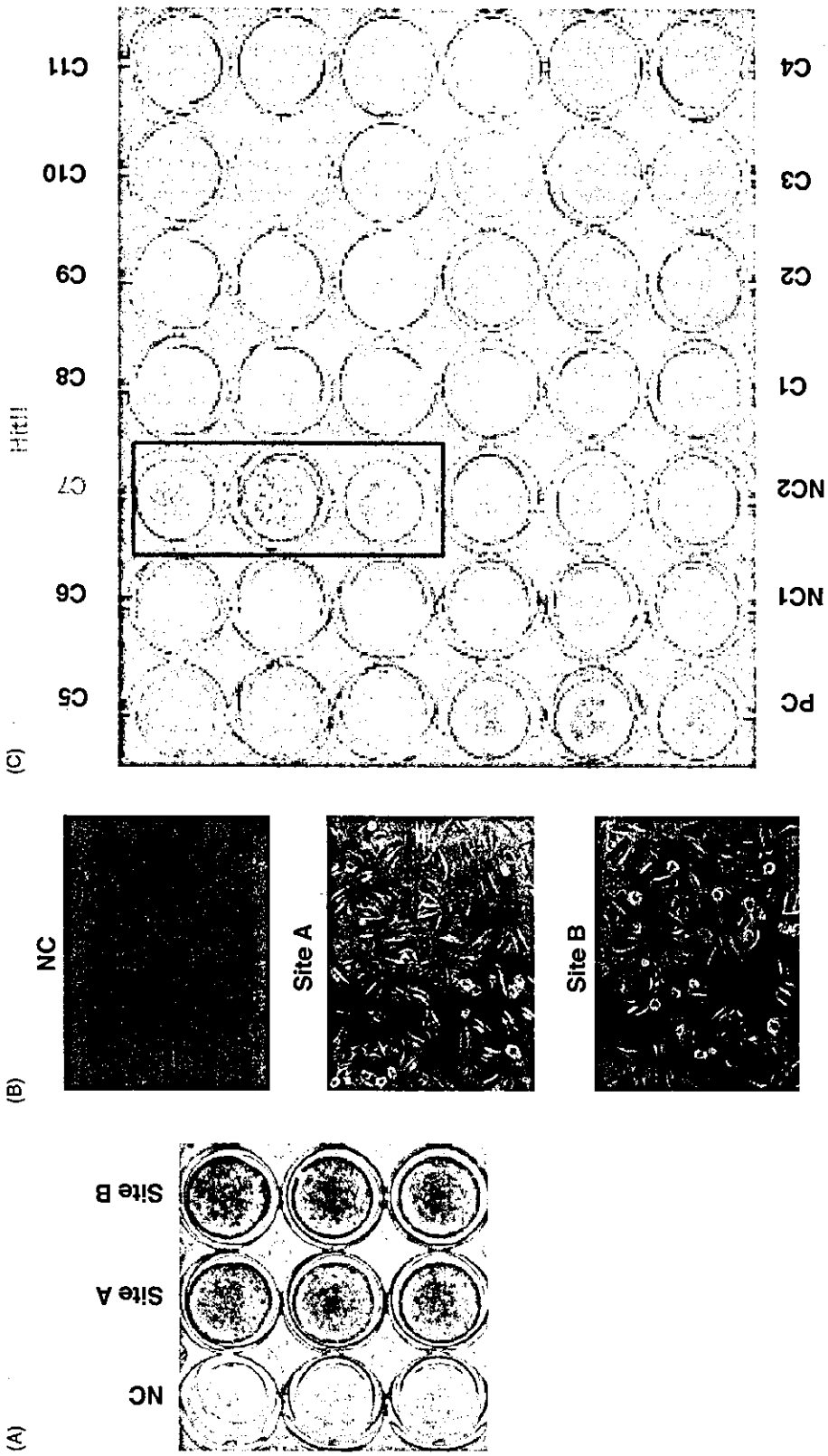

Fig. 5. Identification of genes involved in apoptosis by using shRNAi library. (A) The siRNA-expression vectors targeted against PKR can block dsRNA-induced apoptosis. HeLaS3 cells were transfected with siRNA-expression vectors targeted against two sites of PKR transcript (Site A; 5'TAA TGA ATC AAT CAA TTC ATA TTC-3' and Site B; 5'AAG ACT AAC TGT AAA TTA TGA AC-3'). After the selection with puromycin, the cells were induced to undergo apoptosis by dsRNA transfection. The survived cells were fixed, and stained with crystal violet (0.2%). To check the reproducibility, each independent plasmid was transfected into cells in three different wells (triplicate experiments). NC: negative control (unrelated siRNA-expression vector). (B) Light microscopic images of the cells in the experiment of (B). (C) An example of the screening using siRNA-expression library. PC: positive control; NC1, NC2: negative controls (unrelated siRNA-expression vectors); C1–C11: siRNA-expression vectors targeted against specific genes, for example kinases, transcription factors, or apoptosis-related genes. The shRNA expression vector in C7 inhibits the dsRNA-dependent apoptosis.