

Short hairpin RNA loop design for the facilitation of sequence verification

Oluwaseun Akeju, Tien Peng, and John K. Park
National Institutes of Health, Bethesda, MD, USA

BioTechniques 40:154-158 (February 2006)
doi 10.2144/000112116

The suppression of protein expression using RNA interference (RNAi) has become a powerful tool for the study of gene function (1). Initial applications of RNAi in mammalian cells utilized 21-bp RNA duplexes referred to as small interfering RNAs (siRNAs) (2). Although highly effective, siRNAs were found to be relatively short lived and therefore suboptimal for studies requiring long-term gene suppression. To overcome this limitation, vector expression systems for the DNA directed transcription of single-stranded RNA molecules containing forward and reverse complements separated by a loop segment were devised (3–5). Following transcription, the loop segment enables the formation of a short hairpin RNA (shRNA) structure that can be recognized and processed into siRNAs by the RNase III family nuclease, Dicer (6). Dicer-derived siRNAs can, in turn, be more effective than chemically synthesized siRNAs in inducing messenger RNA (mRNA) degradation (7). A particularly useful aspect of RNAi is the ability to differentially suppress genes or variations of genes that differ by even a single nucleotide. Conversely, an unintended single nucleotide mismatch between an shRNA and its target sequence may be sufficient to abrogate gene suppression (8). It is imperative therefore for shRNA encoding plasmids to be sequence-verified prior to use. The use of standard sequencing methods using DNA polymerases has, however, been problematic because of the inherent secondary structure of the shRNA encoding region. Additives such as dimethyl sulfoxide (DMSO) and betaine, as well as a series of commer-

cially available proprietary compounds, have been used in a trial-and-error fashion to relax the hairpin secondary structure of the single-stranded DNA plasmids during sequencing reactions.

We sought to design a universal loop region that could facilitate the sequence verification of shRNA encoding plasmids. While several algorithms have been developed for the design of the stem portion of an shRNA molecule, such systematic studies have not been reported for the loop region. Given the wide variety of loop region sequences used by individual investigators and the ability of Dicer to efficiently process RNA duplexes lacking loop regions (9), we inferred that the actual nucleotide sequence of the loop is likely subject to minimal constraints. We therefore chose a loop sequence beginning with an “A” followed by the recognition site for the nonpalindromic restriction endonuclease *BmgBI* (5'-A[CACGTC]-3'). Due to its nonpalindromic nature, there is no contribution of this sequence to the stem portion of the shRNA molecule. Likewise, there are no restrictions on the stem sequences adjacent to the loop, as the loop acts as a self-contained unit. We predicted that *BmgBI* digestion of shRNA vectors prior to sequencing would prevent the formation of secondary structures leading to premature termination of DNA polymerase-mediated sequencing reactions.

Single-stranded DNA oligonucleotides were synthesized, annealed, and ligated into the shRNA expression vector pENTR/U6 (Invitrogen, Carlsbad, CA, USA) according to manufacturer's recommendations. Bacteria were transformed with the ligation products and cultured

overnight on antibiotic-containing agar plates. Plasmid DNA was isolated from 10 drug-resistant colonies (QIAprep® Spin Miniprep kit; Qiagen, Valencia, CA, USA) and analyzed by standard dideoxy sequencing using M13 forward and reverse universal sequencing primers and an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster, CA, USA). In the examples presented here, the plasmid clones are intended to encode for shRNA molecules containing 29-bp stem regions directed against stathmin (STMN1, OP18) (Figure 1, A and B). The calculated ΔG of the hairpin structure that can form upon separation of the DNA strands during a sequencing reaction is -48.32 kcal/mol (calculation performed using NetPrimer; Premier Biosoft International, Palo Alto, CA, USA) (Figure 1B). The forward and reverse sequencing reactions of only one of the 10 plasmid DNAs (clone #10) did not terminate within the shRNA encoding region (Figure 1C). The addition of 5% or 10% DMSO to the sequencing reactions yielded qualitatively similar results with 9 of the 10 reactions again terminating prematurely (data not shown). Digestion of the 10 plasmids with *BmgBI* prior to sequencing, however, resulted in successful completion of the sequencing reactions up to the cleavage site (Figure 1D). Of note, clone #10 contains an insertional mutation (denoted in bold type), which likely leads to the formation of less stable hairpin structures during uncut plasmid sequencing reactions (Figure 1E). To underscore the importance of sequence verification of shRNA encoding plasmids, equal DNA quantities of clones #10 and #1 were transfected into U251 malignant glioma cells using FuGENE® 6 (Roche Diagnostics, Alameda, CA, USA). Three days later, whole-cell protein lysates were prepared, separated using polyacrylamide gel electrophoresis (PAGE), and electrotransferred to polyvinylidene difluoride (PVDF) membranes (Immobilon™-P Transfer Membrane; Millipore, Billerica, MA, USA). Membranes were probed with anti-stathmin polyclonal antibodies (Calbiochem, San Diego, CA, USA) and visualized using alkaline phosphatase-linked anti-rabbit

antibodies and dimethylacridinone (DDAO) phosphate (Invitrogen). Membranes were also probed with anti-actin antibodies (Calbiochem) to

confirm equal total protein loading of gel lanes. The mutated shRNA vector is not as effective as the intended shRNA vector in inhibiting the expression of

stathmin. Quantitative analysis of the immunoblot (MultiGauge; Fujifilm, Stamford, CT, USA) reveals a 17% decrease in stathmin expression in clone #10 transfected cells, while clone #1 transfected cells have a 55% decrease (Figure 1F).

The use of DNA directed RNAi is an invaluable tool for the study of gene function. Given the precise nucleotide requirements of the RNAi process, even a single nucleotide mutation within an shRNA molecule can have deleterious effects on gene silencing. Sequence verification of plasmids prior to their use is therefore a necessity. Incorporation of a nonpalindromic restriction site within the loop region of an shRNA encoding plasmid allows for the convenient digestion of the plasmid prior to DNA polymerase-based sequencing. This cleavage of the loop region, in combination with the separate sequencing of the forward and reverse complements that comprise the stem region, significantly increases the likelihood of obtaining reliable full-length sequence information. As demonstrated above, full-length sequencing is not readily achieved unless there is a hairpin destabilizing mutation. Despite a trial-and-error process of optimizing sequencing reactions, correct clones can still sometimes only be identified by the inability to sequence them fully. A potential drawback of our method is that it requires two, both forward and reverse, sequencing reactions rather than one, but each of the two has a high likelihood of success. An additional limitation is the use of a loop nucleotide sequence that may be incompatible with a particular intended siRNA final product. In this event, other nonpalindromic recognition sequences such as those for *Acil* (5'-CCGC-3') or *BsrBI* (5'-CCGCTC-3') can be used. The loop structure and cleavage strategy described above are simple to use and should have broad applicability given the current extensive use of shRNA technologies.

ACKNOWLEDGMENTS

We thank T. Ngo and J. Jarboe for technical assistance and J. Nagle and

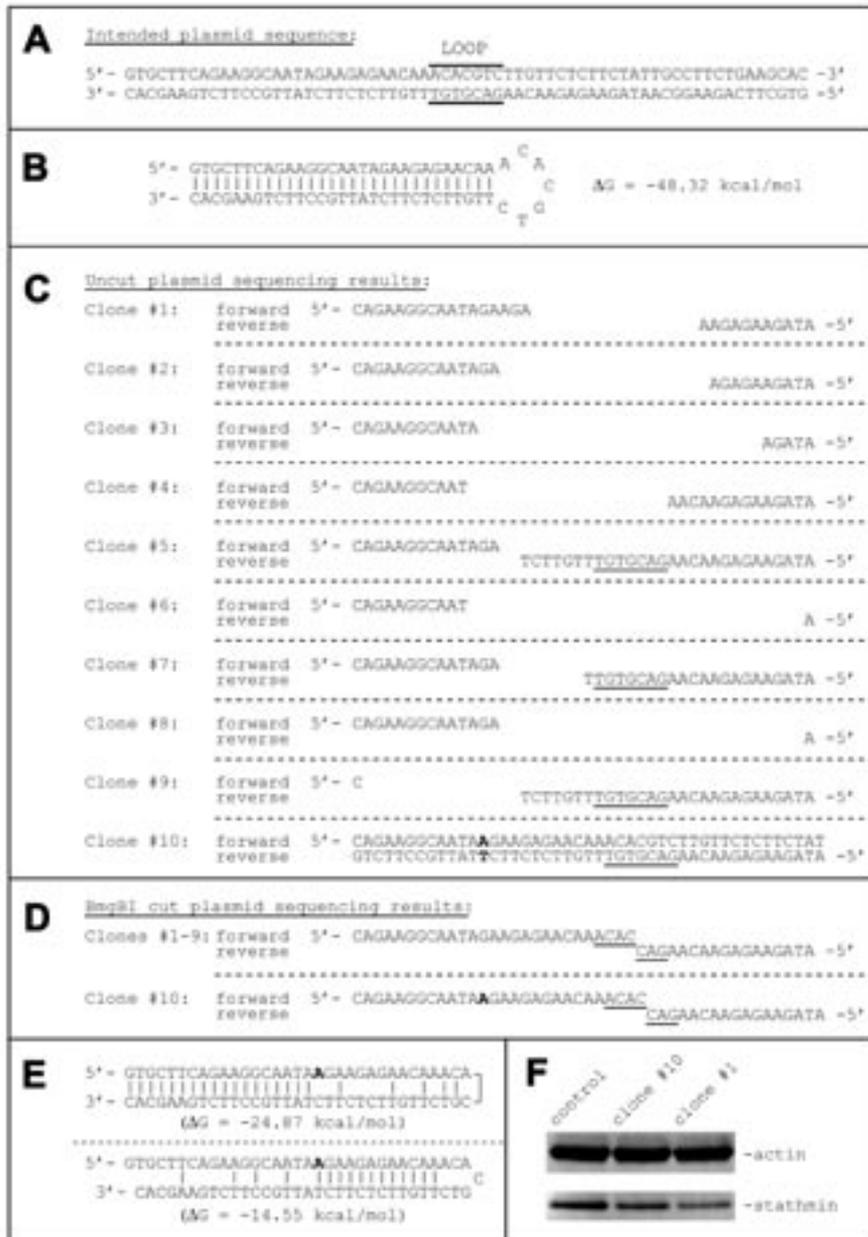


Figure 1. Intended and acquired DNA sequences of short hairpin RNA (shRNA) plasmids. (A) A plasmid for the RNA interference (RNAi) of stathmin expression was designed. Shown is the intended coding region for an shRNA molecule containing a 29-bp stem region and a 7-nucleotide base loop. The loop sequences are underlined. (B) Predicted secondary structure of single-stranded DNA during sequencing reaction. (C) Sequencing results of 10 individual uncut plasmid clones using forward and reverse M13 sequencing primers. The nucleotides 5' to the provided sequences are as predicted and are therefore not shown. Clone #10 appears to have an insertional mutation shown in bold. The loop sequences are underlined. (D) Sequencing results of the same 10 plasmids following digestion with *BmgBI*. Clones #1–9 had identical sequences, and full sequence information up to the cleavage site was obtained. The insertional mutation in clone #10 was again detected. (E) Predicted secondary structures of the two most stable hairpins that can form within clone #10 during a sequencing reaction. (F) Anti-stathmin immunoblot of U251 malignant glioma cells transfected with clone #10 or clone #1 DNA. Anti-actin immunoblot was performed to confirm equal total protein loading of gel lanes.

D. Kauffman of the National Institute of Neurological Disorders and Stroke (NINDS) DNA Sequencing Facility for providing sequencing services. This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), NINDS. T.P. was supported by the Clinical Research Training Program, NIH. O.A. was supported by the Howard Hughes Medical Institute-NIH Research Scholars Program.

Address correspondence to John K. Park, 35 Convent Drive, Room 2B-1002, Bethesda, MD 20892-3706, USA. e-mail: john_park@nih.gov

To purchase reprints
of this article, contact
Reprints@BioTechniques.com

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Hannon, G.J. and J.J. Rossi.** 2004. Unlocking the potential of the human genome with RNA interference. *Nature* *431*:371-378.
2. **Elbashir, S.M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl.** 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* *411*:494-498.
3. **Brummelkamp, T.R., R. Bernards, and R. Agami.** 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* *296*:550-553.
4. **Paul, C.P., P.D. Good, I. Winer, and D.R. Engelke.** 2002. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* *20*:505-508.
5. **Paddison, P.J., A.A. Caudy, E. Bernstein, G.J. Hannon, and D.S. Conklin.** 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* *16*:948-958.
6. **Hammond, S.M., E. Bernstein, D. Beach, and G.J. Hannon.** 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* *404*:293-296.
7. **Siolas, D., C. Lerner, J. Burchard, W. Ge, P.S. Linsley, P.J. Paddison, G.J. Hannon, and M.A. Cleary.** 2005. Synthetic shRNAs as potent RNAi triggers. *Nat. Biotechnol.* *23*:227-231.
8. **Agami, R.** 2002. RNAi and related mechanisms and their potential use for therapy. *Curr. Opin. Chem. Biol.* *6*:829-834.
9. **Kim, D.H., M.A. Behlke, S.D. Rose, M.S. Chang, S. Choi, and J.J. Rossi.** 2005. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* *23*:222-226.

Received 2 November 2005;
accepted 2 December 2005.