mids. This work was supported by the Deutsche Forschungsgemeinschaft (Schwerpunktprogramm: Selenoproteine). Address correspondence to Dr. Marcus Conrad, Institute of Clinical Molecular Biology and Tumour Genetics, GSF Research Centre for Environment and Health, Marchioninistr. 25, D-81377, Munich, Germany. e-mail: marcus.conrad@gsf.de

Received 27 February 2003; accepted 2 April 2003.

Marcus Conrad, Markus Brielmeier, Wolfgang Wurst, and Georg W. Bornkamm
*GSF Research Centre for Environment and Health
Munich, Germany*

Overcoming Obstacles in DNA Sequencing of Expression Plasmids for Short Interfering RNAs

BioTechniques 34:1140-1144 (June 2003)

Gene silencing mediated by RNA interference (RNAi) has been observed in a wide range of eukaryotes including fungi, nematodes, insects, plants, and, more recently, mammalian cells. RNAi has been technologically developed to become a powerful tool to assess the function of genes. RNAi is induced by the presence of dsRNA whose sequence corresponds to a portion of the gene being targeted for silencing. The RNAi mechanism includes the degradation of dsRNA into short interfering RNAs (siRNAs) by Dicer and the degradation of the targeted RNA by RNA-induced silencing complex (RISC) (1–3).

One major barrier to the experimental use of RNAi for targeted gene silencing in mammalian systems was the nonspecific gene silencing triggered by long dsRNAs through the interferon response pathway (4). This difficulty was overcome when Tuschl and colleagues (5) discovered that siRNAs (21 bp), normally generated from long dsRNA during RNAi, can induce a specific and potent RNAi effect in cultured mammalian cells. Initially, siRNAs corresponding to the targeted genes were chemically synthesized *in vitro* and then were transfected into cultured mammalian cells, resulting in specific gene silencing. Soon thereafter, several different plasmids were designed to produce siRNAs *in vivo* from RNA transcripts generated by host cell RNA polymerases (6–11). These plasmids contain a promoter and terminator from genes transcribed by RNA polymerase III (such as the RNase P H1 RNA or U6 snRNA genes) so as to direct the expression of short RNAs with defined ends. Two different designs have been used for these siRNA-expressing vectors. In one strategy, the sense and antisense strands are expressed as two independent transcripts that anneal within the cell to form structures resembling *in vitro* synthesized siRNAs

(7,8). In the second approach, the sense and antisense strands are expressed as a part of a single transcript separated by a short intervening loop. Therefore, the sense and antisense strands form a hairpin structure that after intracellular processing produces siRNAs (6,9,10).

Since sequence complementarity between the siRNA and the targeted gene transcript is critical for efficient RNAi (6,10,12), it becomes important to ensure that the siRNA expression plasmids contain no mutations introduced during the cloning process. However, siRNA vectors that express sense and antisense regions separated by an intervening loop will have hairpin secondary structures in the DNA that can interfere with standard sequencing reactions. Here we describe the use of digestion by restriction enzymes to relieve the secondary structure barriers to allow efficient DNA sequencing of inserts cloned into siRNA vectors.

We employed the pSUPER vector (6) to express siRNAs corresponding to the mammalian transcriptional coactivators BRG1 or p300 by cloning annealed oligonucleotides into the *Bgl*II and *Hind*III sites of pSUPER as described (6). The *Bgl*II site was not restored after cloning. The sequence of the inserted double-stranded oligonucleotide for the BRG1 siRNA is indicated by the capital letters in Figure 1A.

The DNA sequences of the inserted fragments were analyzed using the BigDye™ Terminator Cycle Sequencing Kit on a capillary-based ABI PRISM® 3100 Genetic Analyzer (both from Applied Biosystems, Foster City, CA, USA). Sequencing reactions on the two strands were initiated using primers corresponding to the T7 and T3 RNA polymerase promoters (5'-TAATACGACTCACTATAGGG-3' and 5'-ATTAAACCTCACTAAAGGGA-3', respectively). Problematic sequencing data for both the BRG1 and p300 siRNA plasmids were consistently returned for sequencing reactions from either direction. These reads were typically accurate (Phred values >20) until encountering the inserted DNA oligonucleotides but then prematurely terminated 7–9 nucleotides into the siRNA encoding region (Figure 1B). A likely explanation for the premature termination of the sequencing reactions is the presence of

Benchmarks

strong secondary structure in the DNA insert (Figure 1D) that might inhibit the processivity of the DNA polymerase during DNA sequencing. Attempts to overcome the premature termination by modifying the standard sequencing reactions, including addition of DMSO and increased annealing temperature, were unsuccessful (data not shown).

In an attempt to remove the impeding secondary structure, the siRNA expression plasmids were cleaved with restriction endonucleases that recognize sequences near the loop between the inverted repeats. For example, pSUPER-BRG1 was digested with *BsmBI*, which recognizes the asym-

metric sequence 5'-CGTCTC-3' and cleaves to leave a four-nucleotide 5' overhang. The recognition site for *BsmBI* was fortuitously created near the loop end of the second repeat of the siRNA insert, and the cleavage sites fall within the loop between the inverted repeats (Figure 1A). Initial digestion for 1 h under standard conditions resulted in inadequate cleavage of the plasmid, perhaps because of interference by the DNA secondary structure. Overnight restriction digestion was performed and the resulting DNA subjected to DNA sequencing from both directions. Figure 1C shows that using this strategy the DNA sequencing reac-

tions extended past the site of premature termination to the end of the DNA template at the site of *BsmBI* cleavage. Since the *BsmBI* cleavage site is in the loop region, the sequences of the two inverted repeats were determined from separate reactions using T7 and T3 primers, respectively. The sequences obtained with T3 and T7 primers contain 4 nucleotides of overlap because of the asymmetry in the cleavage by the restriction endonuclease (Figure 1C), ensuring complete sequencing coverage of the insert DNA.

Similar sequencing difficulties were encountered with the undigested siRNA expression plasmid for the p300

Figure 1. Sequence analysis of the pSUPER-BRG1 siRNA expression plasmid. (A) Capital letters indicate the sequence of the DNA insert, and lowercase letters indicate flanking sequence from the pSUPER vector. Letters are colored and aligned to match the sequence chromatograms shown below. Open arrows mark the inverted repeat sequences. A *BsmBI* recognition site (boxed) directs cleavage at sites indicated by arrowheads within the intervening “loop” region. Filled arrows indicate the direction of sequencing reactions using T7 and T3 primers. (B) DNA sequencing chromatograms of the undigested pSUPER-BRG1 plasmid using T7 and T3 primers. (C) DNA sequencing chromatograms of the pSUPER-BRG1 plasmid after digestion with *BsmBI*. (D) DNA secondary structure predicted to occur within the siRNA-encoding region because of sequence complementarity of the inverted repeats. Positions of premature termination of sequencing reactions are indicated with open arrows. Solid arrows indicate run-off termination of sequencing reactions at the template ends following digestion with *BsmBI*.

transcriptional coactivator. In this case, a *Bg*III site had been created at the junction of the loop region with one arm of the inverted repeat. Following digestion with *Bg*III, sequencing reactions proceeded to the ends of the linearized templates, providing a 4-nucleotide overlap in the sequencing products using the T3 and T7 primers (data not shown). This result shows that the sequencing difficulty and the applicability of the solution described herein are not unique to the pSUPER-BRG1 plasmid.

To test whether the release of superhelical torsion of the plasmid pSUPER-BRG1 might be sufficient to obtain successful DNA sequencing, the plasmid was digested using either *Hind*III or *Eco*RI in separate reactions. These enzymes recognize unique restriction sites flanking the DNA insert or H1 RNA promoter, respectively; therefore, digestion with either restriction enzyme resulted in a linear DNA. Nonetheless, the sequencing results following digestion with either *Hind*III or *Eco*RI resembled that of the undigested plasmid, prematurely terminating early in the inverted repeats (data not shown). We conclude that releasing superhelical torsion in the plasmid is not sufficient to overcome the structural barrier presented to the DNA polymerase in DNA sequencing reactions under these conditions.

Our results indicate that siRNA-encoding DNA, with its inherent secondary structure, can be effectively sequenced following restriction digestion to disrupt the secondary structure and separate the inverted repeats. In our example with pSUPER-BRG1, a unique and fortuitous *Bsm*BI restriction site was formed as a consequence of the sequence chosen for its homology to our target gene. In the case of pSUPER-p300, a unique *Bg*III site was formed at the junction of the loop and one arm of the inverted repeat. In neither case did the presence of the restriction sites disrupt the ability of these plasmids to function as RNAi agents. The siRNA expression plasmids targeted to BRG1 and to p300 significantly reduced expression of the target genes in transfected HeLa cells (unpublished observations). Details of that RNAi effect and the consequences for expression of genes at which BRG1 and p300 func-

tion as transcriptional coactivators will be described elsewhere. Other restriction sites could conceivably be similarly engineered within or adjacent to the loop connecting the inverted repeats of siRNA expression plasmids. The restriction sites should be either inherently asymmetric (as for *Bsm*BI in our pSUPER-BRG1 plasmid) or asymmetrically placed (as for *Bg*III in our pSUPER-p300 plasmid) to avoid the potential complications of additional secondary structure inherent in symmetric restriction sites. Moreover, the restriction enzyme chosen should cleave to leave 5'-extended ends so that the DNA sequences coming from opposite directions will overlap.

Several advantages might favor the use of plasmids to generate siRNA over in vitro chemical synthesis, including cost, stability of DNA over RNA, and the possibility of generating stable cell lines expressing a particular siRNA. Since RNAi requires sequence specificity between siRNAs and the targeted RNA, the accurate determination of the DNA sequences cloned into siRNA-producing vectors represents an important step in altering the expression of selected genes using RNAi technology. Although we have demonstrated the value of this tactic using the pSUPER expression vector, the sequencing problem (and thus this solution) are associated with the nature of the inverted-repeat inserts and not to specific components of the vector itself. Thus, this method should be widely applicable for many plasmid or virally based siRNA expression systems.

REFERENCES

1. **Zamore, P.D.** 2002. Ancient pathways programmed by small RNAs. *Science* 296:1265-1269.
2. **Hannon, G.J.** 2002. RNA interference. *Nature* 418:244-251.
3. **McManus, M.T. and P.A. Sharp.** 2002. Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet.* 3:737-747.
4. **Kumar, M. and G.G. Carmichael.** 1998. Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiol. Mol. Biol. Rev.* 62:1415-1434.
5. **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.
6. **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.
7. **Miyagishi, M. and K. Taira.** 2002. U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.* 20:497-500.
8. **Lee, N.S., T. Dohjima, G. Bauer, H. Li, M.J. Li, A. Ehsani, P. Salvaterra, and J. Rossi.** 2002. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* 20:500-505.
9. **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.
10. **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.
11. **Tuschl, T.** 2002. Expanding small RNA interference. *Nat. Biotechnol.* 20:446-448.
12. **Elbashir, S.M., J. Martinez, A. Patkaniowska, W. Lendeckel, and T. Tuschl.** 2001. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* 20:6877-6888.

We thank Dr. R.W. Henry, Dr. Amy Hark, and Dean Shooltz for comments on the manuscript. This work was supported by National Institutes of Health grant no. AI27323 to S.J.T. and by funds from the Department of Biochemistry and Molecular Biology, Michigan State University. Address correspondence to Dr. Steven J. Triezenberg, Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824-1319, USA. e-mail: triezenb@msu.edu

Received 20 February 2003; accepted 3 April 2003.

Daniel C. Ducat, Francisco J. Herrera, and Steven J. Triezenberg
Michigan State University
East Lansing, MI, USA