Complete, gene-specific siRNA libraries: Production and expression in mammalian cells

ATTILA A. SEYHAN, ALEXANDER V. VLASSOV, HEINI ILVES, LEVENTE EGRY, ROGER L. KASPAR, SERGEI A. KAZAKOV, and BRIAN H. JOHNSTON

1SomaGenics, Inc., Santa Cruz, California 95060, USA
2Department of Pediatrics, Stanford University School of Medicine, Stanford, California 94305, USA

ABSTRACT

Short interfering RNAs (siRNAs) are widely used to silence the expression of specific genes. Current practice for designing effective siRNAs is to use algorithms based on sequence-efficacy correlations; however, there are many highly effective sequences that these algorithms do not anticipate. To ensure that the best siRNAs are identified, all possible gene-specific siRNA sequences of appropriate lengths should be screened in cell culture. Synthesizing and testing all such sequences individually is costly. A potentially much easier alternative is to prepare a mixture of all these sequences (a gene-specific library), express them in cells, select cells having the desired phenotype, and identify the siRNA contained within the selected cells. Here we describe two new methods for preparing and expressing such libraries. The first uses cloned Dicer or RNase III to digest gene-specific RNA duplexes to siRNAs, which are then converted to the corresponding DNA sequences by attaching RNA primers and performing reverse transcription-PCR. The second method involves partial DNase I digestion of gene-specific DNA, purification of a 20–30-bp fraction, and amplification by attaching DNA adapters followed by PCR. DNA libraries specific for TNF-α, DsRed, and part of the hepatitis C virus genome, generated by methods, were inserted into siRNA expression vectors between convergent human U6 and H1 promoters. Randomly selected clones from each library together with vectors expressing the corresponding target genes were cotransfected into 293FT cells and assayed for target gene inhibition. About 10%–20% of siRNAs represented in these libraries show significant inhibition of the target genes. Most of these inhibitory sequences are not predicted by existing algorithms.

Keywords: RNA interference; siRNA; pol III promoters; gene inhibition; library; RNAi; Dicer; RNase III

INTRODUCTION

RNA interference (RNAi) is a widespread mechanism in eukaryotes for post-transcriptional gene silencing, associated with the regulation of developmental genes (through micro-RNAs) as well as defense against viral infection and transposon mobility (Hannon and Rossi 2004; Mello and Conte 2004; Meister and Tuschl 2004). It is initiated by the presence of the cell of double-stranded RNA (dsRNA), which is cleaved into ~20-bp fragments by the cytoplasmic enzyme Dicer. Single strands from these fragments become incorporated into an RNA-induced silencing complex (RISC) that can induce cleavage of any mRNA capable of pairing with the RNA strand of the RISC. Long dsRNAs can induce nonspecific gene silencing and apoptosis in mammalian cells, for example by activating the interferon and protein kinase R (PKR) pathways (Leaman et al. 1998; Gil and Esteban 2000; Elbashir et al. 2001a; Hohjoh 2002; Shir and Levitzki 2002). However, introducing 19–29-bp dsRNA fragments (small interfering RNAs, siRNAs) into cells can induce the silencing of specific genes while largely avoiding these nonspecific effects (Zamore et al. 2000; Elbashir et al. 2001a; Scherer and Rossi 2004). Silencing can be induced either by direct transfection of cells with siRNAs or short hairpin (sh) RNAs or by introduction of plasmid or viral vectors expressing siRNAs or shRNAs, usually from pol III promoters (Tuschl 2001; Barton and Medzhitov 2002; Cullen 2002; Devroe and Silver 2002; Hannon 2002; Kitabwala and Ruprecht 2002; Ramaswamy and Slack 2002; Grabarek et al. 2003). In the case of siRNAs, both tandem (unidirectional) (Lee et al. 2002; Lee and Rossi 2004; Miyagishi and Taira 2002) and convergent (opposing) (Tran et al. 2003; Zheng et al. 2004) pol III promoters can be employed for the expression of antisense and sense strands, which then must hybridize to form siRNA duplexes.

Not all siRNA and shRNA sequences are equally potent or specific. Although it had long been thought that siRNAs shorter than ~30 bp avoided induction of interferon and PKR, recent reports indicate that in fact siRNAs longer than...
~19 bp (Fish and Kruithof 2004) or having a 5′-triphosphate group (Kim et al. 2004) can trigger an interferon response. In addition, siRNAs can produce off-target effects, whereby unintended mRNAs are silenced due to having partial homology to the siRNA. Off-target effects may be less problematic with highly potent siRNAs because they can be used at lower concentrations, where discrimination between matched and mismatched targets is greater. Identifying highly potent siRNAs is also crucial to efforts to develop siRNA therapeutics. High potency has been associated with specific sequence features as well as the internal stability profile of the siRNA and the accessibility of the mRNA target site (Elbashir et al. 2001b; Hohjoh 2002; Holen et al. 2002; Lee et al. 2002; Paul et al. 2002; Khvorova et al. 2003; Kretschmer-Kazemi et al. 2003; Reynolds et al. 2004; Ui-Tei et al. 2004). These correlations have been incorporated into algorithms that are commonly used to predict functional siRNAs. Despite their success at finding good siRNAs, many effective siRNA sequences are not predicted by current algorithms. Ideally, all possible target-specific siRNA sequences of appropriate lengths would be tested in cells to assure finding the best inhibitors for a given mRNA (Singer et al. 2004). However, such a “brute force” approach is expensive and time-consuming. An attractive alternative is to screen cell-based libraries of sequences for the most potent siRNAs, without any bias for or against sequence features except for their presence within the target.

In principle, such screening may be performed by using fully random siRNA libraries (Miyagishi and Taira 2002; Tran et al. 2003). However, the high complexity of random libraries (which for 20 random nucleotides is $4^{20} = 10^{12}$) makes this approach impractical for cell-based selection, since there should be one library sequence per cell to make screening straightforward, and $10^{12}$ cells cannot be handled in most laboratories. Alternatively, gene-specific (directed) libraries comprised of all oligonucleotide sequences represented within the target gene (or group of genes) can be used. Their complexity is on the order of the length of the target in nucleotides, which corresponds to a manageable number of cells for screening. Several groups have made gene-specific siRNA pools by digestion of long RNA duplexes with Escherichia coli RNase III (Calegari et al. 2002; Yang et al. 2002, 2004; Kittler et al. 2004) or recombinant human Dicer (Kawasaki et al. 2003; Myers et al. 2003). Such siRNA pools are able to efficiently silence target mRNAs, and can be directly used in cell-based loss-of-function studies, but no selection of the most potent siRNA species is possible unless RNAs are converted into DNA sequences and incorporated into appropriate expression vectors.

Alternatively, siRNA libraries can be generated using gene-specific DNA fragments of appropriate length inserted into expression vectors. We previously described the use of hemi-random probes ligated on a single-stranded DNA target to generate libraries highly enriched in target-specific sequences (Vlassov et al. 2004). Morris et al. (2002) prepared a dsRNA library by inserting 50–500 bp genomic DNA fragments, produced by DNase I digestion of gene DNA, into an expression vector between opposing T7 promoters. However, their method is limited to cells and organisms that are engineered to express T7 RNA polymerase and can tolerate long dsRNA. Recently several groups have generated shRNA libraries by enzymatic digestion of a dsDNA target into fragments and converting the fragments into hairpins of about 20 bp and then into a form suitable for cloning into an expression vector. Shirane et al. (2004) used DNase I to fragment the dsDNA target to an average size of 100–200 bp and blunt-end ligation to attach a hairpin-shaped adapter containing the recognition sequence of Mmel to one end. Mmel cleaves 20 nt to one side of its recognition site; cleavage by this enzyme produced DNA hairpins of an appropriate length. The hairpins were then converted to a double-stranded linear, palindromic form and cloned into an expression vector. Sen et al. (2004) used a somewhat similar procedure except that the dsDNA target was fragmented by using a mixture of frequent-cutting restriction endonucleases instead of DNase I, the fragments were converted into dumbbell-shaped intermediates after Mmel cleavage, and rolling circle amplification was used to generate palindromic sequences for cloning. Both schemes are quite complex. The scheme of Sen et al. is limited by the fact that the cocktail of restriction enzymes does not produce sufficiently random cuts; the reported library contained only 34 unique target-specific sequences out of a theoretically possible 981 for the 1000-nt target. The difficulty of amplifying or transcribing long palindromic sequences and their instability during cloning in E. coli are additional drawbacks that can lead to a reduction in library coverage and potential loss of the best target sites.

An ideal gene-specific siRNA library should have every site represented by multiple overlapping species, should be easy to amplify, clone, and express, and individual species should be 19–29 bp in length. We describe two methods to generate gene-specific libraries of siRNAs that satisfy these criteria and are relatively simple and straightforward. In the first method, double-stranded RNA representing the gene of interest is digested by recombinant human Dicer or bacterial RNase III to produce 20–22 bp siRNAs, which are subsequently converted into corresponding DNA fragments through ligation of adapters (flanking oligoribonucleotides of fixed sequence) and reverse transcription-PCR (RT-PCR). The resulting PCR products are cloned into an siRNA expression vector between opposing human U6 and H1 promoters. In the second method, double-stranded DNA is digested by DNase I in the presence of Mn²⁺ to generate fragments with blunt ends (Melgar and Goldthwait 1968; Campbell and Jackson 1980; Holzmayer et al. 1992). The fraction of short (~20–30 bp) DNA fragments is gel-purified, ligated to adapters, and amplified by PCR. PCR
products of the proper length are again gel-selected and cloned into the siRNA expression vector. Both methods were used to create siRNA libraries directed against the genes for murine TNFα and DsRed. Randomly selected clones from each library were screened for the ability to silence their respective target genes when used with target-expressing plasmids to cotransfect human 293FT cells. This screening showed that ~20% of siRNAs represented in the gene-specific libraries can produce at least moderate inhibition (>60%) of the ectopically expressed target genes. Most of these active species would not have been predicted by existing sequence-based algorithms.

RESULTS

Preparation of gene-specific siRNA libraries by double-stranded RNA fragmentation with Dicer

Two methods for making gene-specific (directed) libraries were developed and compared. In the “Dicer” approach, the dsRNAs corresponding to the cDNA of interest were prepared by transcription from DNA templates flanked by T7 promoters and subsequent annealing. The dsRNAs were then digested with cloned human Dicer as previously described (Kawasaki et al. 2003; Myers et al. 2003) yielding predominantly 21–22-bp siRNAs. Alternatively, modified bacterial RNase III can be used in place of Dicer to generate 20–25-bp siRNA (Yang et al. 2002, 2004). RNase III has the advantages of being less expensive and more efficient than Dicer: at 37°C, RNase III cleavage requires only 20 min for cleavage of a dsRNA substrate as compared to 18 h for Dicer. Both ribonucleases are commercially available (see Materials and Methods). Cleavage products were denatured, purified by denaturing PAGE, and dephosphorylated. RNA adapters were attached sequentially to the 3’- and 5’-ends of the cleavage products by T4 RNA ligase as shown in Figure 1A (Pfeffer et al. 2003). These adapters contain primer-binding sequences and restriction sites. RNAs were subsequently converted into dsDNA fragments by RT-PCR using primers complementary to the adapters. After digestion with appropriate restriction enzymes these gene-specific DNA fragments were ligated into the siRNA expression vector (pU6/H1-coh) (see below and Fig. 3).

Using this approach, libraries specific to three separate genes, TNFα (~1000 bp), DsRed (~840 bp), and the hepatitis C virus internal ribosome entry site (HCV-IRES, ~370 bp), were prepared. To assess the size and distribution along the targets of representative sequences prior to cloning the PCR fragments into the expression vectors, the libraries were cloned into a standard plasmid for sequencing. Sequencing results for randomly chosen clones from the TNF-specific library are shown in Figure 1B. As expected, all inserts were in the range of 21–22 nt and were quite evenly distributed along the target. Of the 27 insert sequences obtained, 24 had a perfect match with target sequences. The other three inserts contained single-nucleotide mismatches or deletions (underlined in Fig. 1B) that were most likely
acquired during the multiple rounds of PCR, which was performed using Taq polymerase. If desired, higher fidelity thermostable polymerases (e.g., Pfu) could be used. However, since such mutations are relatively rare, they do not significantly degrade the quality of the library.

Preparation of gene-specific siRNA libraries by double-stranded DNA fragmentation with DNase I

The second library approach (Fig. 2A) employed partial digestion of the DNA of interest by DNase I in the presence of Mn^{2+} ions, which causes double-stranded cuts and generates blunt-ended fragments (Campbell and Jackson 1980). Reaction times and temperatures were adjusted to maximize the fraction of 20–30-bp digestion products, which were gel-purified by PAGE. These fragments were either directly “blunt-end cloned” into the expression vector or, alternatively, DNA adapters containing PCR primer sequences and restriction sites were attached using T4 DNA ligase, allowing PCR amplification and generation of cohesive ends for convenient insertion into vectors. Using this approach, we generated fragment libraries directed toward two cDNA targets, DsRed (∼840 bp) and HCV-IRES (∼370 bp), and inserted them into the siRNA expression vector (pU6/H1-blunt) used for the Dicer-generated libraries. Sequences of inserts, with flanking pol III transcription terminators, from representative clones of the DsRed-specific library are shown in Figure 2B. All the inserts examined had perfect homology to the targeted gene and were distributed uniformly throughout the DsRed gene. The length of the sequenced inserts varied between 17 and 34 bp, with more than half in the range of 20–30 bp.

These two approaches for generating gene-specific libraries can be considered complementary, with each having certain advantages and disadvantages. The Dicer/RNase III-generated fragments are, of course, the same length as in vivo products of Dicer processing and can be directly incorporated into the RISC complex. The DNase I-generated gene fragments may be more useful for the preparation of shRNA libraries, since the stem length of potent shRNAs can vary from 21 to 29 bp, depending on the sequence (Paddison et al. 2004). Formation of long RNA duplexes from the transcribed antisense and sense strands may sometimes be a challenge for the Dicer/RNase III approach when dealing with highly structured RNAs such as viral internal ribosome entry sites (IRES) elements. On the other hand, the DNase I approach requires at least two gel fractionation steps, and we usually use three (the third after ligation of adapters and PCR).

Gene silencing profiles of randomly selected siRNA clones from the gene-specific libraries

To verify that the U6/H1 expression is working, to check the breadth of coverage of the library, and to roughly estimate the proportion of species that are active at silencing their target mRNA, we cloned, sequenced, and tested in cell culture a number of individual species from libraries generated using each of the methods described. TNFα- and DsRed-specific libraries produced by the Dicer/RNase III and DNase I methods, respectively, were inserted into an opposing-promoter expression vector (Tran et al. 2003; Kaykas and Moon 2004; Zheng et al. 2004). In our version of this vector, called pU6/H1 (Fig. 3), we used human U6 and H1 promoters. We have engineered two versions of this vector to accommodate both cohesive-end and blunt-end cloning of siRNA library inserts (Fig. 3). Individual clones were isolated by growing up single colonies of E. coli transformed by the plasmid libraries. These siRNA clones, together with plasmid vectors expressing

![FIGURE 2. Preparation of gene-specific siRNA libraries by DNase I fragmentation of double-stranded DNA. (A) The general scheme. The double-stranded DNA target is digested by DNase I in the presence of Mn^{2+} ions, and the fraction containing 20–30-bp fragments is gel-purified. Next, double-stranded DNA adapters are attached to 3’ and 5’ ends by T4 DNA ligase, and the resulting fragments are amplified by PCR. Further, fragments are cut with appropriate restriction enzymes and cloned into pU6/H1-coh (see Fig. 3). (B) Sequencing results for the randomly selected clones from the DsRed-specific library.](image-url)
The corresponding target RNAs (TNF or DsRed) and a control target, secreted alkaline phosphatase (SEAP), were used to cotransfect 293FT cells. The vector expressing SEAP was used to control for transfection efficiency and siRNA specificity. Forty-eight hours after transfection, DsRed protein levels were monitored by fluorescence microscopy and quantitated by flow cytometry while TNF protein was assayed by ELISA and SEAP by a colorimetric assay.

The coexpression of TNFα and the individual TNF-specific siRNAs caused down-regulation of TNFα protein synthesis to different extents in different clones (Fig. 4). Of the 36 tested clones, five produced 55%–65% inhibition of secreted TNFα while the others showed <35% inhibition. By comparison, a rationally designed 21-bp siRNA (see Materials and Methods), also expressed from pU6/H1-coh, showed 65% inhibition of TNFα. However, an shRNA positive control having a 29-bp stem and targeting the same region of TNFα mRNA as the 21-bp siRNA control was more potent: When expressed from a plasmid under a uni-directional U6 promoter (pU6), this shRNA inhibited TNFα expression by >95% under similar transfection conditions. As expected, control DsRed-specific shRNA and siRNA expressed from the corresponding vectors had no effect on TNF expression (Fig. 4). The experimental error of individual clones is estimated at ±10%, which was the standard error for triplicate transfections performed with positive control si/shRNAs as well as the sample-to-sample fluctuation in the SEAP negative control.

As with the TNF library, the coexpression of some individual DsRed-specific siRNAs and the DsRed expression vector caused down-regulation of DsRed protein synthesis (Fig. 5).

**DISCUSSION**

The results presented indicate that both types of gene-specific libraries can be expressed from opposing pol III-promoter vectors to produce functional siRNAs. It should be emphasized that the objective of the functional analysis was not to find the most potent inhibitors of gene expression, which requires either a functional screen or much more extensive testing of individual species. Testing of siRNAs randomly selected from gene-specific libraries revealed siRNA species providing a level of gene-specific inhibition similar to that of many siRNAs designed according to current algorithms. Since only limited numbers of the individual library species were tested, it is likely that more extensive testing would reveal even more active sequences. Evaluation of the thermodynamic stabilities and the 5′- and 3′-end sequences of the library siRNAs revealed no correlation between silencing effectiveness and the presence of lower melting, A/U-rich sequences at the 5′-end of the antisense strands, as has been recently reported (Khvorova et al. 2003; Schwarz et al. 2003; Tomari et al. 2004; Hsieh et al. 2004; Ui-Tei et al. 2004). These results suggest that existing algorithms can miss many effective siRNAs; also algorithms cannot predict how variation of length from 21 to 29 bp affects potency or how sequence affects toxicity.

Of 16 clones tested, four showed inhibition of DsRed expression by 55%–70%. As positive controls, a rationally designed DsRed-specific siRNA expressed from pU6/H1-coh and a DsRed shRNA expressed from pU6 showed inhibition of DsRed expression by 81% and 98%, respectively. The negative controls, TNF-specific and eGFP-specific siRNA expressed from pU6/H1, did not inhibit DsRed expression (Fig. 5).
each species and testing them individually in cell culture, a very laborious process (Aza-Blanc et al. 2003; Zheng et al. 2004), or by a screen for the phenotype conferred by inhibition of the target. For fluorescent-tagged targets such as GFP fusions, a fluorescence-activated cell sorter can be used. For targets whose silencing confers a growth or survival advantage, such as a virus or a pro-apoptotic gene, the desired species will outgrow the others. For other targets, fusion with a “suicide gene” such as the thymidine kinase of Herpes simplex virus (HSV-TK) can also allow selection for cells in which the target is silenced (Shirane et al. 2004).

The methods for preparation and expression of gene-specific siRNA libraries described here can also be used to produce libraries of siRNAs targeting a group of genes or an entire genome. Phenotypic screens using such libraries can reveal new potential therapeutic targets as well as new functions of known genes, as has been done using randomized ribozyme libraries instead of siRNAs (Kruger et al. 2000; Li et al. 2000; Kawasaki and Taira 2002; Kawasaki et al. 2002). Any siRNA sequences that may be lethal to the host cell will be eliminated from the library, and sequences toxic enough to retard growth can be eliminated by allowing the cells harboring nontoxic siRNAs to outgrow them.

When provided short targets (<65 bp), Dicer appears to measure from an end in determining its cut sites (Zhang et al. 2002, 2004; Siolas et al. 2004), raising the question of whether sequential cut sites in longer RNAs are in register and might skip over some target sequences. The fact that digestion can occur from either end in most cases provides a second register of cutting that reduces the likelihood of skipping some sequences. Moreover, since each cut site is actually a distribution of several adjacent cleavages (see Zhang et al. 2004), each successive cleavage makes the distribution wider and wider, so that essentially all sites are cleaved except those within about 160–200 bp of the ends. By starting with a dsRNA target flanked by an extra 200 bp of nontarget sequences at either end, this concern can be eliminated, and the resulting addition of a few nontarget siRNAs to the library will have no effect on the effectiveness of library screening. The fact that Dicer cleaves longer dsRNAs more efficiently than shorter ones (Bernstein et al. 2001; Elbashir et al. 2001c; Ketting et al. 2001) suggests that this enzyme may have “endonuclease” activity, independent of ends and therefore not in any fixed register, that is not evident with short fragments where end effects may dominate. Alternatively, fragmentation of a DNA target by DNase I avoids end effects since that enzyme is a true endonuclease. Some sequence preferences can be seen with light digestion (Herrera and Chaires 1994), so adjusting the level of digestion to provide fragments mostly shorter than 30 bp would further reduce the likelihood of missing any sequences in the final library.

In summary, our limited sampling of library sequences together with the above considerations lead to the conclusion that libraries made using Dicer, RNAse III, or DNase I can easily be made to comprise all possible target sequences.

![Figure 4](image_url)

**FIGURE 4.** Silencing ability of species randomly selected from the TNF-specific siRNA library produced by the Dicer method. (A) 293FT cells were cotransfected with randomly chosen clones, a TNF expression vector, and pSEAP using Lipofectamine 2000 (Invitrogen). TNF was assayed by ELISA and SEAP by a colorimetric assay 48 h post-transfection. The inhibition by each siRNA is shown, normalized to the SEAP control target. (B) Representative sequences of the assayed clones were classified into three groups depending on reduction in TNF expression: effective (>60% inhibition), intermediate (20%–60%), and ineffective (<20%).
in the length range of 20–30 nt. The completeness of these libraries, coupled with their cost effectiveness, makes them very attractive alternatives to exhaustive one-by-one testing libraries, coupled with their cost effectiveness, makes them very attractive alternatives to exhaustive one-by-one testing libraries. In a gene walk.

MATERIALS AND METHODS

Preparation of gene-specific libraries by Dicer fragmentation of a dsRNA target

Both sense and antisense strands of RNA target were generated by in vitro transcription from a PCR DNA template having opposing T7 promoters using a MEGAscript T7 kit (Ambion), followed by DNase I treatment, phenol-chloroform extraction, and ethanol precipitation. The RNA strands (1 µg in 10 µL 10 mM Tris-HCl, pH 8.0) were heat denatured at 85°C for 2 min and slowly cooled to room temperature to form dsRNA, which then was digested with 1 unit recombinant Dicer (Stratagene) in 20 µL of 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2.5 mM MgCl₂ for 18 h at 37°C (Fig. 1A). Alternatively, RNA (1 µg) was digested with 1 unit of RNase III (NE BioLabs) in 20 µL of 50 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, 1 mM MnCl₂ for 20 min at 37°C. The resulting 20–22-bp siRNA pool was purified by electrophoresis on a denaturing (7 M urea) 15% polyacrylamide gel. One microgram of the siRNA pool was dephosphorylated with 5 U of calf intestinal alkaline phosphatase (CIAI; MBI Fermentas) for 1 h at 37°C, followed by phenol extraction and ethanol precipitation. All enzymes in this study were used with the buffers provided by the suppliers, as indicated. Next, adapters (single-stranded flanking oligoribonucleotides of defined sequence) were attached to the 3' and 5' ends of each fragment by T4 RNA ligase in two subsequent steps. A typical 40-µL reaction contained 1 µg of fragmented RNA, 200 pmol (~100-fold excess) adapter 1 (5'-p-UUUUUAGAUCGUCCGACAGC-idt-3' [inverted deoxythymidine to prevent circularization]; Dharmacon), 20 U T4 RNA ligase (NE BioLabs), and 40 U RNase OUT (RNase inhibitor; Invitrogen), and the reaction was incubated for 1 h at 37°C. The products were purified by denaturing 15% polyacrylamide gel and ethanol precipitation. RNA (1 µg) was phosphorylated with 20 U of T4 PNK (Polyribonucleotide kinase; NE BioLabs), 1 mM ATP in the presence of 40 U of RNase OUT in a 50 µL reaction for 1 h at 37°C, followed by phenol extraction and ethanol precipitation. The phosphorylated RNA product was ligated to 200 pmol of adapter 2 (5'-GUCGCUGGCAGCUUAAAG-3'; Dharmacon) as described above for adapter 1, followed by phenol extraction and ethanol precipitation. The products of the second ligation were reverse-transcribed using the oligonucleotides attached to the gene-derived sequences as primer-binding sites in a 50-µL reaction containing the RNA product from above, 50 pmol RT primer (5'-GCTGTGGCCAGATCTAAAAA-3'), 20 U AMV-Reverse transcriptase (Promega), dNTPs, and 40 U RNase OUT, 1 h at 42°C. Following reverse transcription, PCR was performed in a 100-µL reaction that contained 10 µL of the reverse transcription reaction, 100 pmol each of the primers 5'-GTGCGTCAGGCTTAAAAA AG-3' and 5'-GCTGTGGCCAGATCTAAAAA-3' (IDT), 5 U Taq polymerase (Promega), 1.5 mM MgCl₂, and 0.2 mM dNTPs (94°C, 30 sec/50°C, 30 sec/72°C, 30 sec, for 10–20 cycles). The ~60-bp PCR products were isolated from a native 10% polyacrylamide gel, cloned into the pT7Blue-3 vector (Novagen), and a number of randomly chosen colonies were sequenced (Retrogen; Fig. 1B). Finally, the PCR products were digested with restriction enzymes and cloned into an siRNA expression vector (see below). Plasmid DNAs isolated (QIAprep Spin Miniprep, Qiagen) from randomly selected clones were sequenced and used for transfection studies.

Preparation of gene-specific libraries by DNase I fragmentation of a dsDNA target

PCR-amplified cDNA encoding DsRed was subjected to partial digestion with DNase I in a buffer containing 1 mM MnCl₂, 50 mM Tris-HCl (pH 7.5), 0.5 µg/µL BSA, and 0.1–0.3 U/µg DNase
I (Ambion) at 20°C for 1–10 min to generate small, blunt-ended DNA fragments (Fig. 2A). Under these conditions DNase I displays little sequence specificity, cleaving all regions of the DNA (except the terminal nucleotides) at an equal rate (Anderson 1981). Since DNase I generates fragments with a wide size distribution, reaction time and temperature were varied to determine optimal conditions to maximize the proportion of DNA in the desired size range (Anderson 1981; Matveeva et al. 1997). Aliquots were collected at various time points and quenched with an equal volume of loading buffer (95% formamide, 10 mM EDTA, 0.1% SDS) and DNA fragments corresponding to 20–30 bp were isolated by native 15% polyacrylamide gel. Next, nicks and potential gaps were repaired by T4 DNA ligase (MBI Fermentas) and DNA pol I (Klenow large fragment; MBI Fermentas) in 50 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 0.1 mM dNTPs, at 20°C for 15 min.

The resulting DNA fragments (which contain 5'-phosphates) can be directly “blunt-end” cloned into the siRNA vector. However, attachment of adapters (fixed flanking double-stranded DNA sequences) is beneficial since it allows PCR amplification and higher ligation efficiency due to the presence of restriction sites in the adapters. The dsDNA adapters were essentially complementary to the 3' termini of modified U6 and H1 promoters (U6: 5'-CTTGTGGAAAGAAGCTTAAAAAG; H1: 5'-AGTTCTGTATGATGACAGATCCTAAAAG). Ligation reactions were performed with T4 DNA ligase, using one adapter at a time, each in ~200-fold excess over the DNA fragments. The ligation products were PCR amplified using primers complementary to the adapter sequences (94°C, 30 sec/52°C, 30 sec/72°C, 60 sec, for 20–30 cycles). The resulting ~70-bp PCR products were purified by native 10% polyacrylamide gel, digested with HindIII and BglII, and, after a second gel-purification, were cloned into the siRNA expression vector (see below). Plasmid DNAs isolated (QiAprep Spin Miniprep; Qiagen) from randomly selected bacterial clones were sequenced and used for transfection studies (Fig. 2B).

Construction of plasmid vectors for siRNA and shRNA expression

To create an siRNA expression vector with convergent pol III promoters, we followed the strategy of Zheng et al. (2004). The promoter regions of the human U6 and H1 small nuclear RNAs were amplified by PCR from genomic DNA of human HT1080 cells by using the following primers: U6 forward (5'-ATCGTCCCGCAGTGGAAAGACGCGCAG); U6 reverse (5'-GATCTCTAACGTCTGTATGATGACAGATCCTAAAAG); H1 forward (5'-ATCGATGCCCCAGTGGAAAGACGCGCAG); H1 reverse (5'-AGTTCTGTATGATGACAGATCCTAAAAG). Ligation of these promoters into the same PCR products containing a loop sequence to facilitate cytoplasmic localization; Lagos-Quin-
produced from covalently closed circular DNA. J. Biol. Chem. 253: 3726–3735.


